



**Titre:** Contributions à la modélisation des déplacements en transport  
Title: collectif

**Auteur:** Mohsen Nazem  
Author:

**Date:** 2014

**Type:** Mémoire ou thèse / Dissertation or Thesis

**Référence:** Nazem, M. (2014). Contributions à la modélisation des déplacements en transport collectif [Thèse de doctorat, École Polytechnique de Montréal]. PolyPublie.  
Citation: <https://publications.polymtl.ca/1539/>

 **Document en libre accès dans PolyPublie**  
Open Access document in PolyPublie

**URL de PolyPublie:** <https://publications.polymtl.ca/1539/>  
PolyPublie URL:

**Directeurs de recherche:** Martin Trépanier, & Catherine Morency  
Advisors:

**Programme:** Génie industriel  
Program:

UNIVERSITÉ DE MONTRÉAL

CONTRIBUTIONS À LA MODÉLISATION DES DÉPLACEMENTS  
EN TRANSPORT COLLECTIF

MOHSEN NAZEM

DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL  
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION  
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR  
(GÉNIE INDUSTRIEL)  
AOÛT 2014

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

CONTRIBUTIONS À LA MODÉLISATION DES DÉPLACEMENTS  
EN TRANSPORT COLLECTIF

présentée par : NAZEM Mohsen

en vue de l'obtention du diplôme de : Philosophiæ Doctor

a été dûment acceptée par le jury d'examen constitué de :

M. SOUMIS François, Ph.D., président

M. TRÉPANIÉ Martin, Ph.D., membre et directeur de recherche

Mme MORENCY Catherine, Ph.D., membre et codirectrice de recherche

M. FAROOQ Bilal, Ph.D., membre

M. HEMILY Brendon, Ph.D., membre

## DÉDICACE

*Nous ne savons pas les secrets éternels, ni toi, ni moi*  
*Nous ne pouvons pas déchiffrer cette énigme, ni toi, ni moi*  
*En deçà du rideau c'est notre discussion*  
*Le rideau écarté, il n'existera ni toi, ni moi !*

*Omar Khayyâm - Mathématicien, philosophe et poète persan du 11<sup>e</sup> siècle*

*À mon cher père, Ibrahim, pour l'amour, le soutien et la patience,*  
*À ma très chère mère, Esmat, pour l'amour, la lumière et la présence,*  
*À mes frères et sœur, Mehrdad, Maziar et Azadeh, pour l'encouragement,*  
*À mes chères belles sœurs, Fatemeh et Masoumeh, pour la présence,*  
*À ma très chère nièce, Newsha, et mon très cher neveu, Shahrokh,*

*À mes grands-parents pour les prières,*  
*À mes oncles et tantes pour l'encouragement,*  
*À mes cousins et cousines pour le soutien,*

*Et à mes très chers amis, pour le soutien moral, pour être ma seconde famille*  
*et pour l'encouragement tout au long de ces années...*

## REMERCIEMENTS

J'ai beaucoup de reconnaissance et d'admiration à témoigner à mon directeur de recherche, le Dr. Martin Trépanier, et ma codirectrice, la Dr. Catherine Morency, qui sont pour moi les modèles parfaits, par leur appui pédagogique, professionnel et personnel.

Également un grand merci pour les chers membres de jury pour l'honneur qu'ils me font en participant à ce jury.

J'adresse un grand merci à mes professeurs tout au long des années d'étude au Canada (École Polytechnique de Montréal), en France (École Nationale des Ponts et Chaussées, École Polytechnique et École Nationale Supérieure des Mines de Paris), en Suisse (Université de Neuchâtel), et en Iran (Université de Technologie Sharif et Université de Technologie d'Isfahan).

Je remercie plus particulièrement les chers membres du département de mathématiques et génie industriel ainsi que du département des génies civil, géologique et des mines de l'École Polytechnique de Montréal.

Je remercie mes chers collègues à l'École Polytechnique de Montréal, surtout les membres de la Chaire Mobilité pour la recherche sur l'évaluation et la mise en œuvre de la durabilité en transport. Un grand merci pour François Godefroy, Audrey Godin, Pierre-Léo Mongeon-Bourbonnais, Félix Pépin, Éric Martel-Poliquin, Louiselle Sioui et Hubert Verreault.

J'adresse un grand merci à tous mes collègues de travail à l'Agence Métropolitaine de Transport, AMT, notamment au sein des Informations Stratégiques et Affaires Métropolitaines, ISAM. Un merci spécial pour Daniel Bergeron, Assumpta Cerda, Alfred Ka Kee Chu, Marie-Christine Desharnais, Charles Lamalice, Kim Mc Donough, Maxime Formond, Nathalie Lacroix, Isabelle Lalouette, Benoît Lavigne, Timothy Spurr et Yonghai Xiao.

Je remercie aussi le Conseil de Recherches en Sciences Naturelles et en Génie du Canada, CRSNG, l'Agence Métropolitaine de Transport, AMT, et d'autres partenaires de mes deux directeurs de recherche pour leur soutien financier indispensable à la réalisation de nos travaux.

Je remercie mes adorables parents, mes chers frères et sœur, mes oncles et tantes, et mes chers cousins et cousines.

J'adresse un grand merci à tous mes amis, pour leurs encouragements et leurs supports moraux. Merci beaucoup Hasan Abdallah, Mani Alizadeh, Glareh Amir Jamshidi, Maher Aoun, Nafiseh Azad, Rayan Bamdad, Mehdi Bavili, Armin Boroumand, Nadim Cherfan,

Amin Dhaou, Mustapha El-Hallak, Farshad Fahim Yahyae, Faranak Fallah, Farsad Farassati, Saeed Farhang, Sahar Ghanipour, Arash Jahangiri, Ahmad Jahanmardi, Sepehr Hamzelounia, Amir Reza Heydari, Farinaz Kalbasi, Hamed Kalantari, Camille Karbassi-Zadeh, Elaheh Khademi, Abdul Kader Khalaf, Elham Khamseashari, Kiana Kiaie, Behnam Maravandi, Saloumeh Massahi, Mohamad Mawlawi, Ali Moayeri, Sina Moetakef, Pouyan Mossanenzadeh, Ali Faryad Nobarian, Emil Osmanovic, Peter Panopalis, Zahra Parvaneh, Sonia Pournazari, Ramin Rad, Farzad Rafieian, Rouzbeh Rahimian, Majid Rasouli, Mohamad Refakar, Ali Rezaie, Banafsheh Reza, Jean Sakr, Rita Sakr, Sahar Samadaei, Helia Serjoui, Amin Aron Shoara, Marc-Antoine Spack, Ciavash Tabari, Morteza Tabatabaei, Ata Taghizadeh, Mehrnoush Talebian, Amir Zahabi, Hoda Zamani, et Minoo Zarabi.

## RÉSUMÉ

Le Transport Collectif (TC) est un composant essentiel de la mobilité urbaine dans le contexte actuel du développement durable. La planification du TC englobe tous les efforts dans les domaines de l'analyse descriptive et de la modélisation. Cette thèse vise à contribuer plus spécifiquement à la modélisation des déplacements en TC.

L'objectif est d'améliorer les modèles en TC pour qu'ils représentent mieux la réalité. Deux types de simplification y sont présents : la construction théorique et la formalisation mathématique. Afin de contribuer à la bonification du processus de la modélisation, cette recherche s'articule globalement autour de trois volets, chacun présenté sous la forme d'un article publié ou soumis pour publication dans une revue scientifique.

Le premier volet consiste au développement d'un modèle de choix d'itinéraires en TC selon les caractéristiques sociodémographiques des utilisateurs. Cela est analysé dans l'article intitulé « Demographic Analysis of Route Choice for Public Transit », et publié dans la revue « Transportation Research Record » en 2011. L'analyse démographique de choix d'itinéraires a permis de segmenter la population en six groupes basés sur les cohortes d'âge et le sexe. L'application d'un modèle de choix discret dans lequel chaque alternative est représentée avec une liste de variables indépendantes décrivant un itinéraire en TC fait partie des contributions essentielles de cet article. Plusieurs segmentations de la population selon les caractéristiques sociodémographiques ainsi qu'un modèle complet considérant la population comme un ensemble monolithique sont examinés. Les résultats obtenus confirment que la segmentation sociodémographique de la population permet d'avoir une meilleure représentation de la réalité en matière de choix d'itinéraires en TC. Certaines bonifications notamment au niveau du raffinement du calcul du temps d'accès en supposant des vitesses de marche différentes selon les caractéristiques sociodémographiques des passagers ont également été présentées.

Le deuxième volet vise le développement d'un modèle d'opportunités pour la génération-distribution des déplacements effectués en TC. Les résultats sont présentés en détails dans l'article intitulé « Integrated Intervening Opportunities Model for Public Transit Trip Generation-Distribution: A Supply-dependent Approach », et publié dans la revue « Transportation Research Record » en 2013. L'avantage principal du modèle d'opportunités pour la distribution des déplacements est qu'il a des bases plutôt comportementales. Nous avons développé et examiné trois formalisations de modèle d'opportunités, nommées modèles fondamental, hiérarchique et

intégré. La comparaison de performances des trois modèles d'opportunités ainsi qu'un modèle gravitaire typiquement utilisé pour la distribution des déplacements confirme une meilleure reproduction de la matrice Origine-Destination (OD) observée à partir d'un modèle intégré d'opportunités. Cet exercice vise, en plus de l'obtention d'un modèle applicable à des fins de prévisions à long terme, à mieux comprendre les phénomènes relatifs à la distribution des déplacements en TC. Sachant que cette distribution est fortement liée à l'accessibilité au TC, celle-ci est prise en compte par le biais des variables dérivées des données opérationnelles d'horaires planifiés et disponibles dans les fichiers General Transit Feed Specifications (GTFS), tout comme certaines variables sociodémographiques et socioéconomiques de la population. Le modèle intégré d'opportunités peut être utilisé en tant qu'un outil de modélisation prospective grâce à la nature prévisible ou supposable de ses variables indépendantes.

Le troisième volet présente un sous-modèle de choix de destination dont bénéficie le développement des modèles d'opportunités. Les constats sont présentés dans l'article intitulé « Revisiting the Destination Ranking Procedure in Development of an Intervening Opportunities Model for Public Transit Trip Distribution », et soumis pour publication dans le « Journal of Geographical Systems » en 2013 ; l'article a été révisé et resoumis en 2014. Ce sous-modèle permet d'améliorer le modèle intégré d'opportunités développé et présenté au deuxième volet sous plusieurs aspects, entre autres la diminution de la sensibilité du modèle aux changements mineurs de temps de parcours en TC. Ce sous-modèle de type de choix discret a permis d'améliorer la performance du modèle d'opportunités, ainsi que d'aider à mieux comprendre le comportement de mobilité de la population utilisatrice du TC.

Les résultats obtenus permettent d'avoir une réflexion plus profonde face à certains enjeux majeurs dont le vieillissement de la population et plus généralement, le changement de structure dans la composition démographique de la population notamment dans les pays développés, et la nécessité d'en prendre considération dans les études de mobilité. Un autre enjeu est d'avoir des outils opérationnels pour examiner à un niveau agrégé, l'impact des projets de transport ou de nouveaux développements urbains sur la génération et la distribution des déplacements. Cette question s'inscrit plus globalement dans l'opérationnalisation des avancées de la communauté scientifique dans la pratique courante des planificateurs de TC. L'émergence de nouvelles données en TC, notamment les données opérationnelles et passives telles que les données GTFS



et de systèmes de paiement par carte à puce (CAP), représentent un atout pour répondre à certaines limites actuelles dans le processus de la modélisation.

En considérant les aspects comportementaux ainsi que l'utilisation des données plus fines, cette thèse contribue à la modélisation en TC sous plusieurs aspects. Les résultats obtenus dans le premier volet pour la considération des caractéristiques sociodémographiques dans le processus de choix d'itinéraires en TC peuvent être intégrés dans les fonctions d'impédance utilisées par les calculateurs de trajet commerciaux comme Google Maps, ainsi que les logiciels de modélisations comme MADIGAS. Concernant le deuxième volet, étant donné des variables utilisées, le modèle intégré de génération-distribution des déplacements peut être appliqué dans les études de prévision de la demande en TC à long terme. Le sous-modèle de choix de destination de nature choix discret développé dans le troisième volet permet d'améliorer la performance du modèle d'opportunités ainsi que d'avoir une meilleure compréhension de la structure de choix de destination des usagers du TC.

La reconsidération des limites observées au cours de cette recherche notamment relatives au niveau de résolution et de disponibilité des données permet d'identifier certaines perspectives pour les recherches futures. Nous proposons entre autre une révision des approches de modélisation aux niveaux théoriques et mathématiques pour les arrimer avec des données nouvellement disponibles en TC.

En raison de la nature appliquée de la modélisation des transports, l'opérationnalisation des modèles est un élément d'intérêt majeur pour la communauté scientifique. Ces modèles permettent de présenter les phénomènes relatifs à la mobilité d'une manière plus réelle pour les professionnels et planificateurs des transports.

## ABSTRACT

Public Transit (PT) is an essential component of urban mobility in the actual context of sustainable development. PT planning contains all efforts for descriptive analysis and modeling. This thesis contributes more specifically to PT modeling.

This thesis objective is to enhance the models in a way that they represent the reality of mobility behavior. Two types of simplifications are present in modeling procedure: Theoretical construction and mathematical formalization. Considering these simplifications, this research contributes in enhancing the modeling procedure in three different ways, each presented as a published or submitted paper in scientific journals.

The first part consists of demographic analysis of route choice for PT. The results are presented in a paper titled “Demographic Analysis of Route Choice for Public Transit”, and published in the journal of Transportation Research Record in 2011. By considering six sociodemographic categories, made up of two genders and three age cohorts, a route choice model for PT is developed and examined. The main contribution of this paper is the development of a discrete choice model for PT route choice based on sociodemographic characteristics. In this model, each choice alternative is represented by a list of independent variables describing a PT route. Several sociodemographic profiles together with a complete model treating the whole population are examined. The results confirmed that considering sociodemographic profiles for analysing PT route choice yields better results compared to the complete model. Some enhancements for calculating access time based on different walk speeds for different sociodemographic profiles are also presented.

The second part is the development of an integrated intervening opportunities model for PT trip generation-distribution. The findings are presented in a paper titled “Integrated Intervening Opportunities Model for Public Transit Trip Generation-Distribution: A Supply-dependent Approach”, which is published in the journal of Transportation Research Record in 2013. The main advantage of the intervening opportunities model is its behavioral basis. In this research, three forms of basic, hierarchical and integrated intervening opportunities model are developed. The performance of this model was analyzed and results were compared with the gravity model typically used for trip distribution. The comparison confirmed that the integrated form of intervening opportunities model could better reproduce an observed OD matrix for PT. Due to sociodemographic, socioeconomic and PT Level-Of-Service (LOS)

independent variables, this model could be used for forecasting purposes. Formalization of this model based on its specific characteristics opens new opportunities for developing more accurate PT demand models with new data such as smart card entries.

The third part discusses a sub-model of destination choice that could be used to enhance the integrated intervening opportunities model developed in the second part. The findings are presented in the paper titled “Revisiting the Destination Ranking Procedure in Development of an Intervening Opportunities Model for Public Transit Trip Distribution”, which was submitted for publication in the *Journal of Geographical Systems* in 2013, and revised in 2014. This sub-model could enhance the integrated intervening opportunities model in different ways, especially for reducing the sensitivity of the model to minor changes in PT trip times. Also, by using this sub-model, the performance of the integrated intervening opportunities model is enhanced. More globally, this discrete choice model could contribute to better understanding of PT travel behavior.

The results obtained in this research could contribute to answer to some challenges facing PT planners and policy makers, including population aging and resulted changes in mobility behavior particularly in developed countries. The PT route choice model based on demographic characteristics developed in the first part could be integrated in commercial PT route choice applications or in simulation software’s impedance functions. Another contribution was presenting an operational tool for examining the impacts of new urban development and transportation projects on PT generation and distribution pattern at high level. The model presented in the second part and enhanced in the third part by using a discrete choice approach could be interpreted in a more general way as a practice-ready approach for PT modeling.

The emergence of new PT data, including operational and passive data such as General Transit Feed Specifications (GTFS) and smart card validations is an opportunity to meet certain limitations in the development of new PT models.

Based on some limitations particularly about the question of data availability, we propose a revision of modeling approaches in both theoretical and mathematical levels to integrate them with newly available PT data.

Given the applied nature of transportation science, the future challenge of the scientific community is developing more practice-ready models. These models could be used by planners and policy-makers for answering to challenges facing PT planning.

## TABLE DES MATIÈRES

DÉDICACE.....	iii
REMERCIEMENTS .....	iv
RÉSUMÉ.....	vi
ABSTRACT .....	ix
TABLE DES MATIÈRES .....	xi
LISTE DES TABLEAUX.....	xv
LISTE DES FIGURES.....	xvii
LISTE DES ACRONYMES ET ABRÉVIATIONS.....	xviii
LISTE DES ANNEXES .....	xx
CHAPITRE 1 : INTRODUCTION.....	1
1.1 Contexte .....	1
1.1.1 Caractérisation de la demande.....	2
1.1.2 Caractérisation de l'offre.....	3
1.1.3 Relation entre la demande et l'offre.....	3
1.2 Objectifs .....	5
1.3 Contributions originales .....	6
1.4 Plan.....	7
CHAPITRE 2 : REVUE DE LITTÉRATURE.....	9
2.1 Approche agrégée.....	12
2.1.1 Procédure séquentielle classique.....	12
2.1.2 Procédure séquentielle classique intégrée .....	25
2.2 Approche désagrégée .....	27
2.3 Approche descriptive.....	28
2.3.1 Sources de données .....	29
2.3.2 Enquête OD, MADITUC et l'approche orientée-objet.....	30
2.3.3 Enquêtes à bord des véhicules.....	33
2.3.4 Potentiel des données opérationnelles et passives.....	34

2.3.5	Systèmes d'aide à la planification .....	39
2.4	Approche de simulation .....	40
2.4.1	Simulation des données de référence .....	40
2.4.2	Simulation des scénarios en transport .....	41
2.5	Approche opérationnelle .....	42
2.6	Synthèse et discussion .....	43
CHAPITRE 3	: DÉMARCHE DE L'ENSEMBLE DU TRAVAIL .....	49
3.1	Considération des caractéristiques sociodémographiques de la population dans la modélisation de leur comportement de choix d'itinéraires en TC .....	49
3.2	Développement d'un modèle intégré de génération-distribution des déplacements en TC à l'aide d'un modèle d'opportunités .....	50
3.2.1	Modèle fondamental d'opportunités (MFO) .....	51
3.2.2	Modèle hiérarchique d'opportunités (MHO) .....	51
3.2.3	Modèle intégré d'opportunités (MIO) .....	52
3.3	Développement d'une procédure de choix de destination de nature choix discret pour l'utilisation dans un modèle d'opportunités .....	53
CHAPITRE 4	: ARTICLE 1 : « DEMOGRAPHIC ANALYSIS OF ROUTE CHOICE FOR PUBLIC TRANSIT » .....	55
4.1	Introduction .....	56
4.2	Background .....	57
4.3	Data set .....	58
4.3.1	Travel Survey Data .....	58
4.3.2	Creation of Alternative Routes Set .....	59
4.4	Development of Route Choice Model .....	60
4.4.1	Definition of Variables .....	63
4.4.2	Descriptive Analysis of Variables .....	64
4.4.3	Estimation of Complete Model for All Passengers as a Whole .....	65
4.4.4	Estimation of Categorical Model for Demographic Groups .....	67

4.4.5	Examples of Route Choice for Demographic Categories .....	69
4.5	Conclusion and Perspectives .....	72
CHAPITRE 5 : ARTICLE 2 : « INTEGRATED INTERVENING OPPORTUNITIES MODEL FOR PUBLIC TRANSIT TRIP GENERATION-DISTRIBUTION, A SUPPLY-DEPENDENT APPROACH » .....		74
5.1	Introduction .....	75
5.2	Literature Review .....	77
5.3	Data for Analysis .....	79
5.3.1	Greater Montreal Area OD Survey .....	79
5.3.2	Census of Canada .....	80
5.3.3	School Enrolment Data Set .....	80
5.3.4	General Transit Feed Specifications .....	80
5.3.5	Data Preparation .....	81
5.4	Development of Integrated Generation-Distribution Model .....	84
5.4.1	Descriptive Analysis of Data .....	85
5.4.2	Development of IIOM .....	85
5.4.3	Modeling Performance Analysis .....	87
5.4.4	Goodness-of-Fit Measures .....	88
5.4.5	Spatial Limitations of IIOM .....	92
5.5	Conclusion and Future Research .....	95
CHAPITRE 6 : ARTICLE 3 : « REVISITING THE DESTINATION RANKING PROCEDURE IN DEVELOPMENT OF AN INTERVENING OPPORTUNITIES MODEL FOR PUBLIC TRANSIT TRIP DISTRIBUTION » .....		97
6.1	Introduction .....	98
6.2	Literature Review .....	101
6.3	Data for Analysis .....	103
6.3.1	Reference Data Sets .....	103
6.3.2	Data Preparation .....	104

6.4	Development of the Models .....	106
6.4.1	General Form of IOM .....	107
6.4.2	BIOM .....	108
6.4.3	EIOM.....	110
6.5	Modeling Performance Analysis .....	114
6.5.1	Cumulative Number of Trips versus Trip Duration .....	115
6.5.2	Trip Production and Trip Attraction Modeling Performance.....	116
6.5.3	Goodness-of-Fit Measures .....	116
6.5.4	Spatial Limitations of the Models .....	118
6.6	Conclusion and Future Research.....	119
CHAPITRE 7 : RÉSULTATS COMPLÉMENTAIRES .....		121
7.1	Raffinement du modèle de choix d'itinéraires en TC en particulierisant la vitesse de marche pour les différents profils sociodémographiques.....	121
7.2	Formalisation d'un modèle d'opportunités pour la distribution des déplacements tous motifs en TC.....	126
CHAPITRE 8 : DISCUSSION GÉNÉRALE .....		128
CHAPITRE 9 : CONCLUSION .....		133
9.1	Synthèse des travaux .....	133
9.2	Contributions.....	135
9.3	Limitations .....	135
9.4	Perspectives.....	136
BIBLIOGRAPHIE .....		139
ANNEXES .....		156

## LISTE DES TABLEAUX

Tableau 2.1 Résumé des points forts et faibles des approches de modélisation en transport .....	46
Tableau 2.2 Résumé des points forts et faibles relatifs aux données en transport .....	47
Tableau 4.1 t-Test for All Mode and Transfer-Type Variables and for All Demographic Categories.....	61
Tableau 4.2 t-Test for All In-Vehicle and Walking Times and for All Demographic Categories.	62
Tableau 4.3 Sample of Integrated Data Set Used for Analysis .....	64
Tableau 4.4 Descriptive Analysis of Variables for Two Data Sets.....	66
Tableau 4.5 Estimation of Utility Function Coefficients by Gender .....	67
Tableau 4.6 Estimation of Utility Function Coefficients by Age Cohort .....	69
Tableau 4.7 Estimation of Utility Function Coefficients by Age Cohort and Gender .....	70
Tableau 5.1 Sample of Number of Spatially Cumulative Opportunities for Work Trips .....	84
Tableau 5.2 IOM Calibrated Parameters .....	87
Tableau 5.3 Goodness-of-Fit Measures .....	90
Tableau 6.1 Sample of Spatially Cumulative Opportunities for Work Trips (Basic Destination Ranking Procedure).....	110
Tableau 6.2 Sample of OD Trip Numbers Derived from the 2008 GMA OD Survey .....	113
Tableau 6.3 Sample of Data Layout for Conditional Logit Model (Enhanced Destination Ranking Procedure) .....	113
Tableau 6.4 Calibrated Coefficients of the Utility Function (Enhanced Destination Ranking Procedure) .....	114
Tableau 6.5 Sample of New Destination Ranks Based on Probabilities (Enhanced Destination Ranking Procedure).....	114
Tableau 6.6 Calibrated Parameters of BIOM and EIOM.....	115
Tableau 6.7 Similarity Measures for Comparing Modeled Cumulative Number of Trips versus Trip Duration.....	115
Tableau 6.8 Goodness-of-Fit Measures .....	118
Tableau 7.1 Écart entre les valeurs observées et celles modélisées pour les itinéraires non-reproduits par le modèle complet (vitesse de marche constante).....	122
Tableau 7.2 Vitesse de marche selon les profils sociodémographiques .....	123



Tableau 7.3 Coefficients de la fonction d'utilité du modèle de choix d'itinéraires en TC par sexe (vitesses de marche différentes) .....	124
Tableau 7.4 Coefficients de la fonction d'utilité du modèle de choix d'itinéraires en TC par cohortes d'âge (vitesses de marche différentes).....	124
Tableau 7.5 Coefficients de la fonction d'utilité du modèle de choix d'itinéraires en TC par cohortes d'âge et par sexe (vitesses de marche différentes).....	125
Tableau 7.6 Comparaison de performances des modèles de choix d'itinéraires en TC.....	125
Tableau 7.7 Paramètres de MIO (déplacements tous motifs) .....	127
Tableau 7.8 Mesures statistiques de validité de l'ajustement des MFO, MHO, MIO et modèle gravitaire (déplacements tous motifs) .....	127
Tableau A.1 Synthèse chronologique de la revue de littérature.....	156

## LISTE DES FIGURES

Figure 1.1 Procédure générale simplifiée de prise de décision dans le domaine des transports.....	4
Figure 2.1 Schéma simplifié de la PSC.....	11
Figure 2.2 Schéma simplifié d'un modèle d'activités.....	11
Figure 2.3 Schéma simplifié de la PSC avec rétroaction .....	13
Figure 2.4 Méthodologie de la production des données pour l'analyse et la modélisation en transport.....	16
Figure 2.5 Schéma de la POTC en trois niveaux d'étude .....	44
Figure 4.1 Examples of Route Choice in the Greater Montreal Area .....	71
Figure 5.1 Spatial Distribution of the MS on Island of Montreal .....	77
Figure 5.2 Data Flow Diagram for Development of IIOM .....	79
Figure 5.3 Deriving the Reference OD Matrices and Trip Production and Attraction Values from the OD Survey .....	82
Figure 5.4 Cumulative Number of Trips versus Trip Duration.....	89
Figure 5.5 Spatial Residual Errors Reported from IIOM.....	93
Figure 6.1 Spatial Distribution of MS in the Island of Montreal .....	100
Figure 6.2 Data Flow Diagram for Development of BIOM and EIOM.....	105
Figure 6.3 Cumulative Number of Trips vs. Trip Duration for BIOM and EIOM .....	116

## LISTE DES ACRONYMES ET ABRÉVIATIONS

AFC	Automatic Fare Collection
AMT	Agence Métropolitaine de Transport
AOT	Autorité Organisatrice de Transport
APC	Automatic Passenger Counting
AVL	Automatic Vehicle Location
BIOM	Basic Intervening Opportunities Model
CAP	Carte à Puce
CASPT	Conference on Advanced Systems for Public Transit
CATS	Chicago Area Transportation Study
CBD	Central Business District
CRSNG	Conseil de Recherches en Sciences Naturelles et en Génie du Canada
CTM	Commission de Transport de Montréal
DI	Dissimilarity Index
EIOM	Enhanced Intervening Opportunities Model
ESRI	Environmental Systems Research Institute
FAMOS	Florida Activity Mobility Simulator
GIS	Geographical Information System
GM	Gravity Model
GMA	Greater Montreal Area
GTFS	General Transit Feed Specifications
HAGS	Household Attributes Generation System
IIOM	Integrated Intervening Opportunities Model
IMPACT	Integrated Model for Population Ageing Consequences on Transportation
IOM	Intervening Opportunities Model
ISAM	Informations Stratégiques et Affaires Métropolitaines
LOS	Level-Of-Service
MADGEN	Modèle d'Analyse Désagrégée des Générateurs de Déplacements
MADITUC	Modèle d'Analyse Désagrégée des Itinéraires de Transport Urbain Collectif
MAE	Mean Absolute Error

MFO	Modèle Fondamental d’Opportunités
MG	Modèle Gravitaire
MHO	Modèle Hiérarchique d’Opportunités
MIO	Modèle Intégré d’Opportunités
MS	Municipal Sector
NMAE	Normalized Mean Absolute Error
NPHI	Normalized PHI
NPTS	Nation-wide Personal Transportation Survey
NSERC	Natural Sciences and Engineering Research Council of Canada
OD	Origine-Destination
ODS	Origin Destination Survey
PDA	Personal Digital Assistant
PME	Percentage Misallocated Error
POTC	Planification Opérationnelle en Transport Collectif
PSC	Procédure Séquentielle Classique
PT	Public Transit
PTLOS	Public Transit Level-Of-Service
RMSE	Root Mean Square Error
RP	Revealed Preferences
SIG	Systèmes d’Informations Géographiques
SM	Secteur Municipal
SP	Stated Preferences
STL	Société de Transport de Laval
STM	Société de Transport de Montréal
STO	Société de Transport de l’Outaouais
TAZ	Transportation Analysis Zone
TC	Transport Collectif
TRB	Transportation Research Board
VBA	Visual Basic for Applications

## **LISTE DES ANNEXES**

Annexe A	Synthèse de la revue de littérature .....	156
----------	---	-----

## **CHAPITRE 1 : INTRODUCTION**

Avez-vous pris le transport collectif ce matin pour vous rendre au travail ? Êtes-vous maintenant assis sur un banc du métro en train de lire ? Ou peut-être vous êtes en train de lire dans un abribus en attendant le prochain passage...

Comme il m'a dit un inconnu à Paris en 2007, on ne peut pas comprendre une ville sans connaître son transport !

### **1.1 Contexte**

Le Transport Collectif (TC) est un composant majeur de la gestion urbaine dans le contexte actuel du développement durable. Il s'inscrit dans l'objectif des modes de transport durables et de la préservation de l'environnement, tout en apportant des avantages sociaux et économiques pour l'ensemble de la population.

La mobilité constitue un élément essentiel de la qualité de vie urbaine, puisque la majorité des citoyens sont bénéficiaires des services de transport (Fedra, 2003). Le transport est un problème multidisciplinaire qui est lié étroitement aux politiques de l'utilisation du sol, aux enjeux énergétiques, à la structure économique de la société et à la qualité de l'environnement. Vu sa complexité et sa diversité, aucune approche d'étude unique n'arrive à couvrir toute la gamme des questionnements. La planification du TC peut être divisée en deux catégories d'étude distinctes mais fortement liées entre elles et généralement d'une manière rétroactive : l'analyse descriptive et la modélisation.

La première catégorie consiste à étudier l'état actuel ou l'état prévu des phénomènes relatifs à la mobilité en TC. Par exemple, les études sur le comportement de mobilité des usagers du TC font partie de cette catégorie. La deuxième catégorie englobe tous les efforts déployés afin de produire des modèles en TC, y compris les études destinées à la prévision de la demande en TC à différents horizons. Dans cette recherche, nous nous attarderons plus particulièrement à la problématique de la modélisation des déplacements en TC.

Selon l'encyclopédie Larousse, un modèle dans le domaine scientifique est une « *représentation abstraite des relations entre les paramètres caractéristiques d'un phénomène ou d'un processus. Un modèle scientifique vise à simplifier le phénomène étudié, à unifier des phénomènes a priori sans rapport et à prédire des phénomènes nouveaux* » (Larousse, 2014). Plus précisément, un modèle mathématique est « *une traduction de la réalité pour pouvoir lui appliquer les outils, les techniques et les théories mathématiques, puis généralement en sens inverse, la traduction des résultats mathématiques obtenus en prédictions ou opérations dans le monde réel* » (Caussidier, 2007).

La problématique de modélisation en TC est très diverse et complexe. Bien qu'elle ne soit pas nouvelle, nous voulons y contribuer en nous appuyant d'une part sur les aspects comportementaux de la mobilité en TC, et d'autre part sur les enjeux informationnels relatifs aux données requises pour des fins de modélisation. La recherche s'articule sur l'amélioration de la modélisation du TC. Pour ce faire, nous devons tout d'abord mieux connaître les différentes approches de la modélisation, classées ici comme agrégée, désagrégée, descriptive, de simulation, et opérationnelle. Cette compréhension nous permet de déterminer les aspects comportementaux de chaque approche ainsi que la structure de données nécessaires pour chacune d'entre elles. L'objectif de cette thèse est de contribuer à la modélisation en TC, plus particulièrement aux approches agrégée, désagrégée et descriptive.

Afin de mieux apprécier les enjeux relatifs à la modélisation en TC, nous distinguons les problèmes en transport en trois catégories tirées d'une vision économique, et expliquées dans les sections suivantes (Leurent, 2007; Ortuzar & Willumsen, 1994).

### **1.1.1 Caractérisation de la demande**

Une approche simple dans le domaine de la planification du TC consiste à ne pas du tout estimer la demande ! Dans ce cas, nous sommes dans un monde totalement expérimental (CUTS, 2010). Cela représente une approche très risquée. En effet dans la plupart des cas, en raison des enjeux majeurs associés aux projets de transport, la prévision de la demande est inévitable. Les enjeux économiques associés aux investissements nécessaires, les délais de construction et les longues durées de vie de certains infrastructures et équipements, les relations entre le transport et les développements économiques et sociaux ainsi que les impacts environnementaux sont parmi les raisons qui justifient la nécessité de faire des études de planification de TC (Bonnel, 2004).

### **1.1.2 Caractérisation de l'offre**

L'offre de TC est un service et non pas un bien. Pour bénéficier de ce service, il faut le consommer où et quand il est produit. Autrement dit, on ne peut pas stocker de service de transport, d'où l'importance d'avoir une adéquation entre la prévision de la demande et l'offre en TC (Ortuzar & Willumsen, 1994).

### **1.1.3 Relation entre la demande et l'offre**

La demande de TC est dépendante d'une part des caractéristiques sociodémographiques et socioéconomiques de la population concernée, et d'autre part des caractéristiques de l'offre de transport (Webster & Bly, 1982). La modélisation de transport contribue à la connaissance de ces éléments, et ultérieurement conseille les fournisseurs de TC en matière de développement et d'amélioration des services offerts.

Il existe de multiples enjeux relatifs au processus général de la modélisation. Certains chercheurs ont présenté les lacunes majeures en lien avec des résultats obtenus des modèles de prévision et de leur transférabilité dans le temps et l'espace. Ils citent des problèmes associés au processus de collecte de données et à la modélisation, ainsi que leurs effets cumulatifs (Atkins, 1976; Flyvbjerg, Holm, & Buhl, 2006). Malgré toutes les limitations mentionnées, il nous semble que la planification du TC par le biais des analyses descriptives et de la modélisation est une approche scientifiquement crédible. Par contre, il reste une longue route à faire pour faire face aux nouveaux enjeux et problématiques qui émergent chaque jour.

Selon Manheim, il existe deux défis essentiels (Manheim, 1979) :

- Le défi de substance de l'analyse systémique des transports est d'intervenir délicatement et délibérément sur le tissu complexe d'une communauté afin d'utiliser le transport efficacement, en concordance avec d'autres actions publiques ou privées, dans l'atteinte des buts de cette communauté.
- Le défi méthodologique de l'analyse systémique des transports est de mener pour une situation particulière, une analyse systématique qui soit valide, pratique et pertinente, et qui contribue à la clarification des enjeux à débattre.



La planification des transports dans ses deux aspects d'analyse descriptive et de modélisation forme un outil d'aide à la décision pour répondre aux enjeux relatifs à la mobilité.

La procédure générale simplifiée de prise de décision dans le domaine de transport par la voie de planification est présentée dans la Figure 1.1 (Bonnell, 2004). Face à un problème, le développement de solutions alternatives et l'évaluation de leurs performances pour en choisir la meilleure peuvent être également ajoutés dans la procédure (Beimborn, 2006).

Deux types de simplification sont présents dans la procédure de la modélisation des transports : la construction théorique et la formalisation mathématique. Le premier consiste à choisir des variables selon les relations causales entre elles et à étendre ces relations logiques dans le futur. Quant au deuxième, il repose sur l'utilisation des modèles mathématiques qui impliquent nécessairement des simplifications dues à différentes raisons, généralement le manque de données nécessaires (Bonnell, 2004).

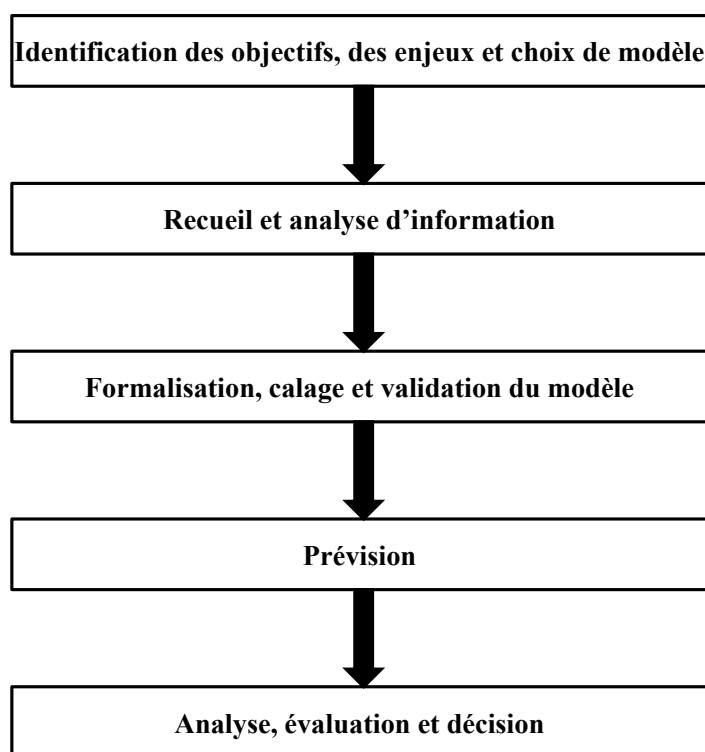


Figure 1.1 Procédure générale simplifiée de prise de décision dans le domaine des transports  
(inspirée de Bonnell, 2004)

Dans cette recherche, nous visons la bonification de la planification du TC, tout en contribuant à la modélisation en proposant de nouvelles approches de choix de variables et de calage des modèles ainsi que de nouveaux ensembles de données. Le but final est de présenter les phénomènes réels relatifs à la mobilité en TC par la voie de la modélisation mathématique.

## 1.2 Objectifs

L'objectif général de cette recherche est de contribuer à l'analyse descriptive et à la modélisation du TC. Les perspectives de cette recherche s'inscrivent dans une approche améliorée en planification du TC.

Notre objectif spécifique est d'améliorer la construction théorique et la formalisation mathématique de certains modèles en TC.

Plusieurs enjeux et pistes de réflexion sont pris en considération au cours de cette recherche. En voici quelques-uns :

- Le vieillissement de la population et plus généralement le changement de structure de la composition démographique de la population sont des phénomènes dans plusieurs pays développés (StatCan, 2010). Cela nous amène à l'idée de prendre en considération certaines caractéristiques sociodémographiques de la population utilisatrice du TC dans la modélisation de leur comportement de mobilité sous plusieurs aspects.
- Les caractéristiques socioéconomiques de la population ont des effets sur le comportement de mobilité des personnes, d'où l'importance de les apprécier dans le processus de la modélisation.
- La prévision de la demande de TC en fonction des changements apportés à l'offre a une grande importance pour assurer un équilibre acceptable entre la demande et l'offre. Elle s'inscrit dans la perspective de multiples études en transport, parmi lesquelles citons les études d'ensemencement pour estimer la demande générée lors de nouveaux projets en transport. Dans cette optique et en lien avec les points précédents, nous développons un modèle intégré d'opportunités pour la génération-distribution des déplacements en TC en utilisant des variables sociodémographiques, socioéconomiques et de l'offre de TC.
- L'émergence des nouvelles données fines telles que les données d'horaires planifiés provenant de General Transit Feed Specifications (GTFS), de validation des cartes à puce (CAP), de localisation des véhicules du TC provenant de systèmes Global Positioning

System (GPS) ouvre de nouvelles portes pour examiner de multiples variables, ainsi qu'un calage plus précis des modèles en TC.

- L'opérationnalisation des modèles développés par la communauté scientifique pour les rendre applicables aux cas réels auxquels font face les planificateurs de TC est un enjeu essentiel. Cela est associé plus particulièrement au choix de variables, à la disponibilité de données requises et à l'applicabilité technique des modèles au cas divers.

Nos travaux permettent d'avoir une réflexion plus profonde face à plusieurs problématiques dans le domaine de la modélisation du TC ; parmi celles-ci citons :

- Comment considérer les caractéristiques sociodémographiques et socioéconomiques de la population dans les modèles de prévisions de déplacements en TC, notamment les modèles de choix d'itinéraires ?
- Comment estimer la demande en fonction des changements apportés à l'offre de TC à un niveau agrégé surtout pour des fins de prévisions de mobilité à long terme ?
- Comment analyser le choix de destination des usagers du TC à l'aide d'une approche d'analyse comportementale, plus précisément un modèle de choix discret au sein d'un modèle d'opportunités pour la distribution des déplacements ?

La section suivante fait ressortir les contributions originales de cette recherche.

### **1.3 Contributions originales**

Cette recherche contribue à l'amélioration des approches de modélisation en TC. Cela se fait en considérant les aspects comportementaux et à l'aide des données plus fines pour contribuer à une meilleure représentation de la réalité par la voie de la modélisation mathématique sous plusieurs aspects :

- La considération des caractéristiques sociodémographiques dans la procédure de choix d'itinéraire dans un réseau de TC. Les résultats peuvent a posteriori être intégrés dans les fonctions d'impédance utilisées par les logiciels de modélisation tels que MADIGAS, etc.
- Le développement d'un modèle d'opportunités sensible à l'offre de TC ainsi qu'aux caractéristiques sociodémographiques et socioéconomiques de la population pour la

génération-distribution des déplacements en TC. Cela représente un modèle intégré pour reproduire la matrice Origine-Destination (OD) en TC à partir des variables indépendantes prévisibles ou supposables pour le futur. Ce modèle est pertinemment applicable en pratique, notamment dans les études de prévisions de mobilité à long terme attribuables aux projets majeurs de TC, telles que l'étude d'ensemencement.

- Le développement d'un sous-modèle de choix de destination conceptualisé dans le but d'apporter des améliorations au modèle d'opportunités. Ce sous-modèle de nature choix discret permet d'améliorer la performance du modèle d'opportunités mentionné ci-dessus, mais également de mieux comprendre la structure de choix de destination des usagers du TC.

La section suivante est dédiée à la présentation du plan de la thèse.

## 1.4 Plan

Cette thèse présente les contributions apportées à la modélisation des déplacements en TC sous plusieurs volets. Le Chapitre 1 est dédié à l'introduction, composé du contexte, des objectifs et des contributions originales de la recherche.

Le Chapitre 2 présente une revue de littérature pertinente aux travaux. Celle-ci s'attaque premièrement aux fondements des méthodes de planification du TC, tels que l'approche agrégée présentée par la Procédure Séquentielle Classique (PSC), où l'on cite notamment la problématique des modèles de distribution. À ce sujet, le modèle d'opportunités qui représente une nouvelle façon de traiter la distribution des déplacements est expliqué. L'approche désagrégée est discutée brièvement par la suite. La revue s'attarde ensuite sur l'approche descriptive qui est basée sur l'analyse des résultats obtenus en présence de données nombreuses comme c'est le cas dans la grande région de Montréal (Québec). Ensuite, le Modèle d'Analyse Désagrégée des Itinéraires de Transport Urbain Collectif (MADITUC), et l'approche orientée-objet sont décrits, suivis des travaux touchant l'utilisation des données opérationnelles et passives dont l'information provient des systèmes de paiement par CAP. Les deux sections suivantes de la revue sont dédiées aux approches de simulation et opérationnelle en TC. Enfin, une synthèse des avantages et des limites de chaque approche de la modélisation et des données relatives est présentée.

Le Chapitre 3 présente la démarche de l'ensemble du travail, ou l'organisation générale de la thèse pour atteindre ses objectifs en lien avec les limites présentées à la fin du Chapitre 2. Globalement les travaux de cette thèse sont divisés en trois volets, présentés en détails aux Chapitres 4, 5 et 6 respectivement :

- la considération des caractéristiques sociodémographiques des voyageurs dans la modélisation de leur comportement de choix d'itinéraires en TC ;
- le développement d'un modèle intégré de génération-distribution des déplacements en TC à l'aide d'un modèle d'opportunités sensible aux caractéristiques sociodémographiques et socioéconomiques de la population ainsi qu'à l'offre de TC ;
- le développement d'une procédure de choix de destination améliorée pour l'utilisation dans un modèle d'opportunités pour la génération-distribution des déplacements en TC.

Chacun de ces trois chapitres contient un article publié ou soumis pour publication dans les revues scientifiques.

Dans le Chapitre 7, quelques résultats complémentaires sont présentés sous deux sections. La première consiste à présenter des améliorations apportées aux résultats obtenus au cours du Chapitre 4. Quant à la deuxième section, nous présentons trois différents modèles d'opportunités pour les déplacements de tous motifs en TC : modèles fondamental, hiérarchique et intégré. Les deux premiers modèles sont discutés en détails dans un article présenté lors d'une conférence (Nazem, Trépanier, & Morency, 2012), et le modèle intégré est expliqué dans le Chapitre 7.

Le Chapitre 8 présente une discussion générale, en rappelant les objectifs et contributions de la thèse dans le domaine de la modélisation en TC. Dans ce chapitre, quelques pistes de développement qui nécessitent plus d'approfondissement méthodologique et technique sont également présentées.

Le Chapitre 9 est dédié à la conclusion et quelques pistes de recherches pour le futur.

La thèse se termine par une liste bibliographique ainsi qu'une annexe qui présente un tableau synthétique décrivant la revue de littérature par une liste chronologique des chercheurs et de leur sujet de recherche respectif.

Dans cette thèse, pour assurer l'homogénéité de rédaction tout au long du document peu importe la langue de rédaction, les valeurs décimales sont présentées avec « , ».

## CHAPITRE 2 : REVUE DE LITTÉRATURE

Bien qu'il existe différentes approches en modélisation des transports, vu l'objectif de cette recherche qui est la contribution à la modélisation des déplacements en TC, la revue de littérature est ciblée sur les approches appliquées particulièrement au TC. Nous présentons d'abord un préambule à cette revue pour déterminer les approches les plus couramment utilisées.

La modélisation du TC joue un rôle indispensable dans la procédure de la prise de décision face aux problématiques de la mobilité. Vu la diversité et la multitude des approches présentées dans le domaine, la première étape consiste à choisir l'approche la plus appropriée dans chaque cas d'étude. Le choix d'approche est généralement fait à partir de certains éléments tels que le contexte du problème, la précision nécessaire, la disponibilité des données requises, l'état de l'art de la modélisation et la disponibilité des ressources économiques, techniques et humaines (Ortuzar & Willumsen, 1994).

La littérature plus récente identifie deux approches en modélisation de transport : l'approche basée sur les déplacements et l'approche basée sur les activités (McNally & Rindt, 2007).

La première approche, qui englobe diverses familles de modèles classiques basés sur les déplacements a toujours omis de reproduire clairement le comportement des voyageurs. La Procédure Séquentielle Classique (PSC) fait partie de cette famille de modèles. Malgré le manque de représentation de comportements des voyageurs dans certains cas, il était évident dès le début que la nature dérivée de la demande de transport a été comprise et acceptée, mais pas forcément reflétée dans la formalisation des modèles. Ces modèles se sont développés dans les années de boom de l'après-guerre et ont été utilisés et modifiés jusqu'à aujourd'hui. Cependant, dans les années 70, les changements fondamentaux dans les zones urbaines, les enjeux économiques, énergétiques et environnementaux ainsi que les avancées technologiques ont influencé les approches de modélisation en transport. C'est dans cette période que les modèles basés sur les activités ont été lancés et étudiés (McNally & Rindt, 2007).

La Figure 2.1 présente un schéma simplifié des quatre étapes d'un modèle de type PSC, faisant partie des approches basées sur les déplacements, et la Figure 2.2 résume les étapes d'un modèle de type activités.

Les modèles basés sur les déplacements de type PSC et ceux d'activités sont bâtis selon les philosophies différentes. La PSC modélise les déplacements au niveau zonal et les modèle d'activités simulent les activités des personnes et des ménages en considérant les chaînes de

déplacements. Une étude en Floride a démontré une comparaison des résultats obtenus à partir de ces deux approches de modélisation, en comparant avec le comportement observé (Shan, Zhong, Du, & Lu, 2012). Bien que cette étude démontre certaines forces et faiblesses de chaque approche, nous ne recommandons pas de généraliser les résultats obtenus, car ils peuvent être dépendants de comportement de mobilité observé dans le contexte de l'étude de cas réalisée.

Dans chacune de ces deux approches, plusieurs familles de modèles sont développées afin de permettre la réalisation des étapes différentes. Nous visons à étudier différentes approches et modèles examinés au cours des années, pour mieux comprendre leur fonctionnement ainsi que leurs limitations.

À partir de la littérature concernée, nous distinguons cinq approches différentes dans les études du TC. Les sections suivantes présentent les principales études et expérimentations associées à chaque approche.

L'objectif de cette revue de littérature n'est pas de faire une description exhaustive des modèles de prévision de la demande en transport. Ce qu'on cherche est davantage une compréhension des limites des modèles au niveau de la construction théorique et de la formalisation mathématique. À partir des limites constatées, nous visons par la suite à améliorer la procédure de modélisation des déplacements en TC afin de mieux représenter les phénomènes réels.

Certaines approches étudiées dans ce chapitre sont basées sur les déplacements et d'autres sur les horaires planifiés. Cette revue de littérature est composée de cinq sections : l'approche agrégée, l'approche désagrégée, l'approche descriptive, l'approche de simulation, et l'approche opérationnelle dans la modélisation des déplacements en TC.

Au cours de cette revue, certaines approches en lien avec les analyses présentées aux Chapitres 4, 5, 6 et 7 sont présentées en détails, alors que d'autres approches sont abordées de manière générale.

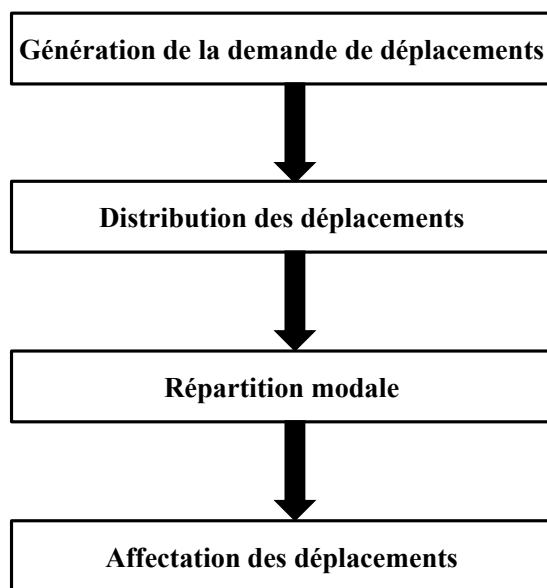


Figure 2.1 Schéma simplifié de la PSC

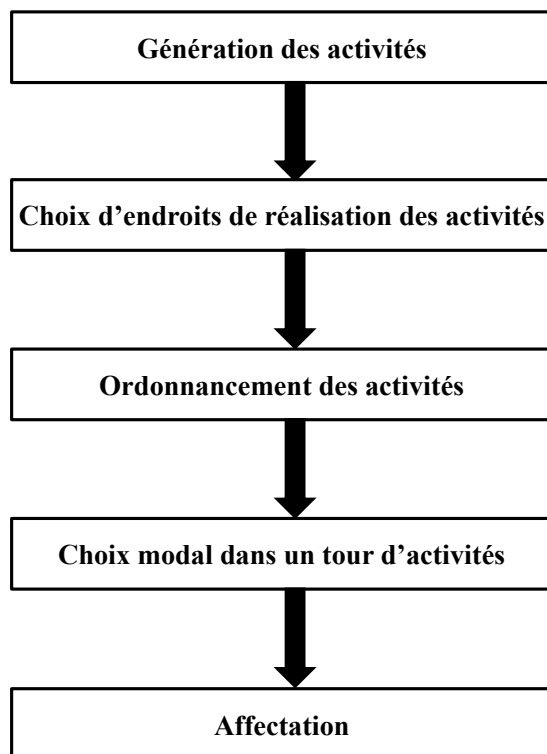


Figure 2.2 Schéma simplifié d'un modèle d'activités  
(inspiré de Roorda et Miller, 2006)



## 2.1 Approche agrégée

L'approche agrégée en planification de transport est l'approche dans laquelle l'analyse est généralement basée sur un système zonal. Autrement dit, en appliquant cette approche, on obtient des résultats sur une base zonale et non pas individuelle. Dans cette section, les modèles agrégés en transport sont abordés. Nous essayons lorsque possible, de cibler la revue sur les modèles développés en TC, mais dans certains cas, des explications sur la procédure générale de la modélisation sont indispensables. Sachant que la plupart des modèles sont dérivés de l'idée de la PSC, nous présentons tout d'abord cette procédure. Pour ce faire, le développement de divers modèles mathématiques pour la réalisation des étapes de la PSC, soit la génération, la distribution, la répartition modale et l'affectation des déplacements est discuté.

Par la suite, deux approches d'intégration dans le cadre de la modélisation agrégée sont présentées. La première vise l'intégration de certaines étapes de la PSC afin de les réaliser d'une manière simultanée et non pas séquentielle ; la deuxième consiste à intégrer la PSC avec des modèles démographiques ou environnementaux pour prendre en considération de façon plus explicite les effets sociodémographiques dans la modélisation en transport.

### 2.1.1 Procédure séquentielle classique

La PSC représente l'approche agrégée classique de la modélisation en transport. Elle est composée de quatre étapes dont la génération, la distribution, la répartition modale, et l'affectation des déplacements (Bonnell, 2004).

Les deux approches agrégées principales sont la PSC sans rétroaction (Figure 2.1) et la PSC avec rétroaction (Figure 2.3). Ce qui distingue ces deux approches est que la PSC avec rétroaction représente d'abord la réalisation des quatre étapes d'une manière séquentielle et ensuite la modification des étapes réalisées précédemment à partir des flux estimés sur le réseau. Par contre, dans la PSC sans rétroaction, après avoir estimé des flux sur le réseau, il n'y a pas de retour sur les étapes réalisées précédemment.

La notion de rétroaction peut être appliquée de façons différentes. Dans une étude récente de Boyce et al. trois procédures de rétroaction sont appliquées sur un modèle de prévisions de déplacements à Albany (New York) (Boyce, O'Neill , & Scherr, 2008) :

- rétroaction directe dans laquelle, après avoir réalisé l'étape de l'affectation, les matrices de déplacements et les flux sont réaffectés directement sur le réseau. Cette procédure est

continué jusqu'à ce qu'on arrive à une certaine convergence. Autrement dit, aucune moyenne sur les matrices de déplacements, ni sur les flux n'est appliquée.

- rétroaction dans laquelle on fait la moyenne sur les matrices de déplacements avec des poids constants.
- rétroaction dans laquelle on fait les moyennes successives avec des poids différents sur les matrices de déplacements.

Les résultats de l'application de ces approches à la région d'Albany (New York) montrent que la procédure de rétroaction directe n'est pas efficace. Les auteurs trouvent également que l'application de poids constants est plus efficace que la moyenne successive (Boyce et al., 2008). Bien évidemment, les résultats obtenus ne sont pas globaux et sont sensibles à chaque cas. Pour arriver aux conclusions plus globales, il faudra tester la procédure de la rétroaction dans des études de cas différentes.

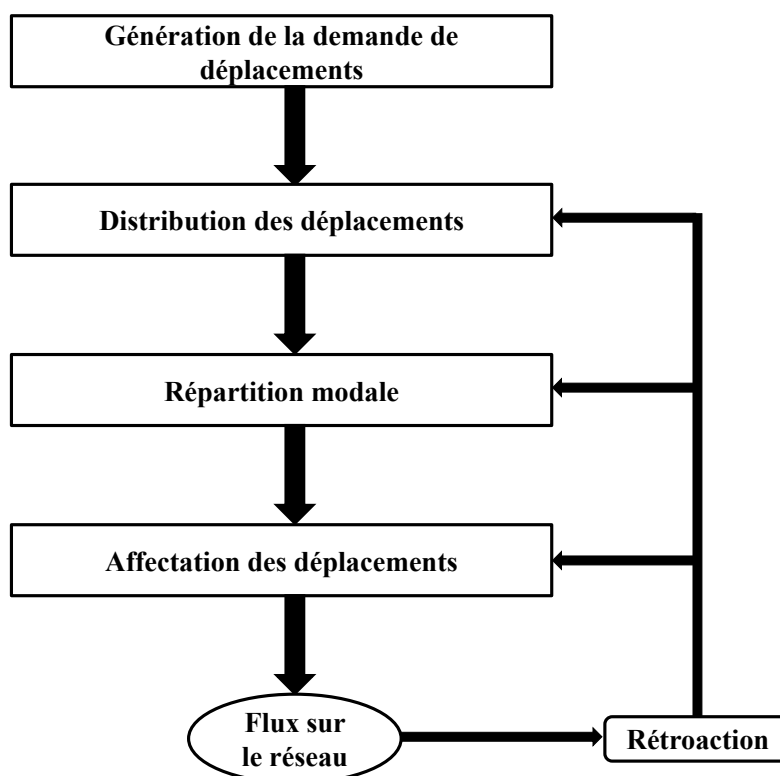


Figure 2.3 Schéma simplifié de la PSC avec rétroaction  
(inspiré de Hensher et Button, 2008)

Avant de réaliser une étude de la PSC, certaines données dont les caractéristiques sociodémographiques de la population, l'aménagement du territoire, les services de transport et les statistiques concernant l'usage de ces services sont requises (Rajesson, 2009). La Figure 2.4 démontre la méthodologie de production et préparation des données pour des fins de modélisation (Bonnell, 2004).

Nous présentons ici très brièvement les quatre étapes de la PSC pour la prévision de la demande en transport (Bonnell, 2004; Hensher & Button, 2008; Meyer & Miller, 2001; Ortuzar & Willumsen, 1994).

### 2.1.1.1 Génération

La génération de la demande en transport consiste à déterminer le nombre de déplacements entrant et sortant pour chaque zone d'étude. Dans la plupart des cas, les déplacements sont estimés par motif, dont le travail, l'étude, le magasinage, le loisir, etc. Les principales familles de modèles de génération sont les modèles normatifs, les modèles de régression linéaire et les modèles de classification croisée (Bonnell, 2004). Nous présentons brièvement les modèles de régression linéaire multiples.

#### 2.1.1.1.1 Modèles de régression linéaire multiple

Le modèle le plus couramment utilisé est celui de la régression linéaire multiple. L'idée principale est que les déplacements sont produits et attirés en fonction des paramètres de l'utilisation du sol et de la population. La forme globale d'un modèle de régression linéaire est :

$$Y = \sum_{i=1}^n \beta_i X_i + \epsilon \quad (2.1)$$

$Y$  : Taux ou nombre de déplacements

$n$  : Nombre de variables indépendantes utilisées dans le modèle

$\beta_i$  : Coefficients de la régression linéaire

$X_i$  : Variables indépendantes

$\epsilon$  : Erreur

Dans la littérature, certaines variables indépendantes sont utilisées pour la production des déplacements dont la population, le revenu, l'état de motorisation et la composition du ménage. En ce qui concerne l'attraction, les variables utilisées dépendent des motifs de déplacement ; par

exemple le nombre d'emplois peut représenter l'attraction d'une zone pour les déplacements motif travail, et le nombre de places scolaires dans une zone pour modéliser l'attraction quant aux déplacements motif étude (Bonnell, 2004).

Lorsqu'il y a une grande variété comportementale, la segmentation des ménages ou des individus devient intéressante. Suite à la segmentation, on pourra appliquer le modèle de régression linéaire à chaque segment.

Les plus grandes lacunes des modèles de génération sont (Bonnell, 2004) :

- la formalisation linéaire des modèles, qui est une hypothèse a priori, pas forcément représentative de la réalité ;
- la multi colinéarité entre variables indépendantes ;
- l'hypothèse d'homoscédasticité en supposant une variance constante pour les nombres de déplacements prédits dans les zones différentes, et la nécessité de garder une certaine homogénéité dans le découpage zonal.

Cette méthodologie est utilisée de façon générale pour réaliser l'étape de la génération des déplacements. Nous supposons que l'application de cette méthodologie pour déterminer le nombre de déplacements générés et attirés en TC soit possible avec certaines modifications dans le choix de variables indépendantes.

Après avoir réalisé l'étape de la génération, il faudra modéliser la distribution des déplacements, la deuxième étape de la PSC.

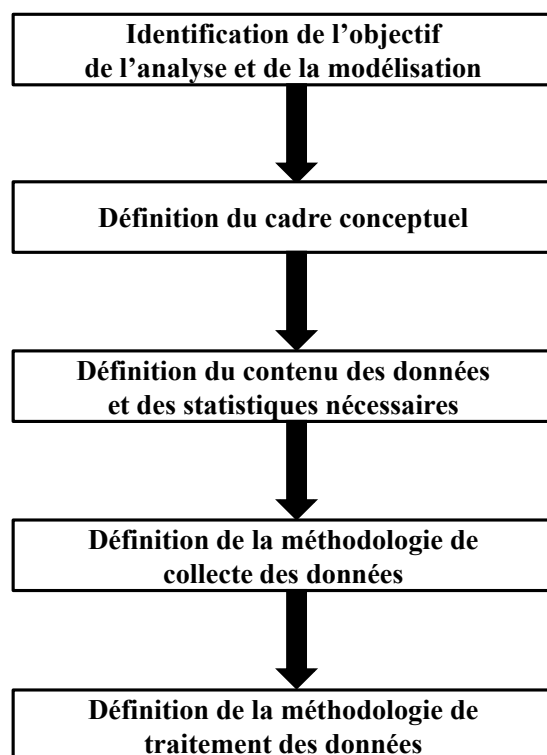


Figure 2.4 Méthodologie de la production des données pour l'analyse et la modélisation en transport

(inspirée de Bonnel, 2004)

### 2.1.1.2 Distribution

L'objectif de la deuxième étape de la PSC est de déterminer une matrice OD des déplacements. Dans cette étape, nous cherchons à modéliser le nombre de déplacements par paire OD dans le territoire d'étude. Trois familles de modèles sont présentées dans la littérature : les modèles basés sur les facteurs de croissance, le modèle gravitaire et le modèle d'opportunités (Bonnel, 2004).

D'autres modèles mathématiques sont également expérimentés, dont un modèle basé sur le concept des destinations compétitives. Dans ce modèle, un problème d'optimisation multi-objectif est formalisé en utilisant les notions d'accessibilité et d'attractivité pour des paires OD. Le modèle gravitaire présenté ci-après peut être également dérivé de cette formalisation mathématique (De Grange, Troncoso, Ibeas, & González, 2009).

#### 2.1.1.2.1 Modèles à facteurs de croissance

Le modèle à facteurs de croissance consiste à transformer une matrice de référence, soit à partir d'un facteur uniforme, soit à partir des facteurs pour chaque ligne et chaque colonne. À première vue, cela représente une approche simple. Par contre, il est nécessaire d'avoir une matrice de référence, ce qui n'est pas toujours le cas. De plus, en utilisant cette approche il devient plus difficile de tenir compte d'une nouvelle offre de transport (TC ou autre). Autrement dit, cette approche ne permet pas de prendre en considération l'impact de l'évolution de l'offre du TC sur le patron de déplacements (Bonnell, 2004).

#### 2.1.1.2.2 Modèle gravitaire

Le modèle gravitaire qui s'inspire de la loi de Newton, utilise une fonction d'impédance face aux déplacements qui est souvent définie à partir de distance ou plus généralement de coût généralisé pour chaque paire OD : plus la zone de destination est éloignée de la zone d'origine, moins l'utilisateur est enclin à s'y rendre. Sa formalisation classique est la suivante :

$$T_{ij} = A_i E_i B_j D_j f(c_{ij}) \quad (2.2)$$

$T_{ij}$  : Nombre de déplacements entre les zones  $i$  et  $j$

$A_i$  et  $B_j$  : Coefficients interdépendants mis à jour à chaque itération qui traduisent le respect des contraintes aux marges telles que  $A_i = 1 / \sum_{j=1}^n B_j D_j f(c_{ij})$  et  $B_j = 1 / \sum_{i=1}^n A_i D_i f(c_{ij})$

$E_i$  et  $D_j$  : Nombre total des déplacements produits (émis) par la zone  $i$  et attirés (se destinant) par la zone  $j$

$f(c_{ij})$  : Coût généralisé de déplacement entre les zones  $i$  et  $j$

La notion  $f(c_{ij})$  représente la fonction d'impédance. Dans le cas du TC, cette fonction peut avoir des composants différents de l'automobile tels que le temps d'attente, les temps d'accès à l'origine et à la destination, le nombre de correspondances, le nombre de lignes empruntées, le coût monétaire, etc.

Après avoir formalisé le modèle gravitaire, il est important d'examiner sa performance. Les mesures statistiques utilisées à cet effet sont celles de la validité de l'ajustement

(Goodness-of-Fit) et celles de la comparaison macroscopique à l'aide des mesures comparatives basées sur la distance et le temps de parcours (Thamizh Arasan, Wermuth, & Srinivas, 1996).

La calibration d'un modèle de distribution se fait de façons différentes : si on a accès à  $E_i$  et  $D_j$  d'une manière indépendante du modèle de distribution, les deux équations suivantes servent à satisfaire des contraintes sur les déplacements produits et attirés, respectivement (Wilson, 1970) :

$$\sum_j T_{ij} = E_i \quad (2.3)$$

$$\sum_i T_{ij} = D_j \quad (2.4)$$

Les équations suivantes donnent les variables  $A_i$  et  $B_j$  qui peuvent être utilisées dans le modèle gravitaire :

$$A_i = \frac{1}{\sum_j B_j D_j f(c_{ij})} \quad (2.5)$$

$$B_j = \frac{1}{\sum_i A_i E_i f(c_{ij})} \quad (2.6)$$

Les Équations 2.3 et 2.4 représentent les contraintes sur la production et l'attraction des déplacements.

Supposons maintenant que nous avons accès uniquement à  $E_i$ . Dans ce cas-là, on a une seule contrainte qui est sur les déplacements produits et le modèle prend la forme suivante, où  $W_j^{(2)}$  présente une notion d'attractivité de la zone  $j$  (Wilson, 1970) :

$$T_{ij} = A_i E_i W_j^{(2)} f(c_{ij}) \quad (2.7)$$

$$A_i = \frac{1}{\sum_j W_j^{(2)} f(c_{ij})} \quad (2.8)$$

On peut présenter la notion  $T_{*j}$  comme suit :

$$T_{*j} = \sum_i T_{ij} \Rightarrow T_{*j} = W_j^{(2)} \sum_i A_i E_i f(c_{ij}) \quad (2.9)$$

Puisque la destination d'un déplacement est un indice de l'activité, cela présente une forme de modèle de localisation des activités.

Dans une autre situation, on peut avoir des contraintes uniquement sur l'attraction des déplacements. Dans ces modèles en effet, nous n'avons accès qu'à  $D_j$ . La variable  $W_i^{(1)}$  représente un indice de production pour la zone  $i$  (Wilson, 1970) :

$$T_{ij} = W_i^{(1)} B_j D_j f(c_{ij}) \quad (2.10)$$

$$B_j = \frac{1}{\sum_i W_i^{(1)} f(c_{ij})} \quad (2.11)$$

et également :

$$T_{i*} = \sum_j T_{ij} \Rightarrow T_{i*} = W_i^{(1)} \sum_j B_j D_j f(c_{ij}) \quad (2.12)$$

Finalement, on présente la famille des modèles qui ne présentent pas de contraintes, ni sur les déplacements produits,  $E_i$ , ni sur ceux attirés,  $D_j$  :

$$T_{ij} = K W_i^{(1)} W_j^{(2)} f(c_{ij}) \quad (2.13)$$



Dans ce dernier modèle, nous utilisons des facteurs de production et d'attraction au lieu des valeurs de production et attraction des déplacements,  $E_i$  et  $D_j$ . Dans la réalité, on utilise rarement cette dernière formalisation, car nous avons généralement accès à au moins une des deux variables de  $E_i$  ou  $D_j$  (Wilson, 1970). Après avoir calibré le modèle, il faudra valider les résultats obtenus (Thamizh Arasan et al., 1996).

La calibration des modèles de ce type est en général faite à partir de certaines mesures telles que  $R^2$  pour identifier l'écart entre les valeurs estimées et celles de référence. Cette approche rencontre parfois des problèmes d'endogénéité (De Grange et al., 2009). Par la définition, le problème d'endogénéité dans un modèle se produit quand il y a une corrélation entre les variables indépendantes et les termes d'erreur (Kennedy, 2003). Une façon d'éviter ce problème est d'utiliser des variables proxy pour la production et l'attraction des déplacements ainsi qu'une procédure d'optimisation. Par contre, il a été démontré que les erreurs survenues à cause d'un problème d'endogénéité n'ont pas d'effets significatifs lors de l'utilisation d'un modèle de type gravitaire (De Grange et al., 2009).

Néanmoins, de façon générale, le modèle gravitaire a toujours été remis en question en raison de ses faibles bases comportementales.

#### 2.1.1.2.3 *Modèle d'opportunités*

Le modèle d'opportunités a des bases plutôt comportementales par rapport au modèle gravitaire en se basant sur la probabilité de choisir une destination à partir de l'origine (Kermanshah, 2004). L'idée fondamentale du modèle est que la distance ou plus généralement le coût généralisé n'est pas l'élément essentiel face à la réalisation d'un déplacement. Ce modèle considère en fait l'accessibilité relative des opportunités qui satisfont le motif de déplacement. L'idée s'inscrit dans les fondements théoriques qui relient la mobilité et la distance à la migration et à la localisation des services (Stouffer, 1940, 1960) :

*“The number of persons going a given distance is directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities” (Stouffer, 1940).*

Schneider a développé la théorie du modèle d'opportunités dans la façon présentée aujourd'hui (Schneider, 1959). Dans le domaine du transport, ce modèle est brièvement présenté dans la littérature (Bonnel, 2004; Hensher, 1977; Ortuzar & Willumsen, 1994).

Selon le modèle d'opportunités, un individu choisit la destination la plus proche de son lieu d'origine, qui lui offre l'opportunité de satisfaire son besoin. Dans ce modèle, la distance n'est pas une variable continue telle qu'elle est utilisée dans le modèle gravitaire ; elle sert plutôt à trouver l'ordre des zones à partir d'une origine donnée, autrement dit l'ordonnancement des zones de destination (Bonnel, 2004).

Pour appliquer ce modèle, il faut énumérer des zones séparément pour chaque origine et ensuite, les classer par ordre croissant de distance ou de coût généralisé. La probabilité de choisir une zone par un individu est égale à la probabilité de ne pas choisir les zones plus proches à l'origine, multipliée par la probabilité de trouver une opportunité satisfaisante pour le motif de déplacement dans cette dernière zone. La forme générale d'un modèle simple-contrainte est (Bonnel, 2004) :

$$T_{ij} = E_i \cdot \frac{e^{-pO_{k-1,i}} - e^{-pO_{k,i}}}{1 - e^{-pO_N}} \quad (2.14)$$

$O_{k,i}$  : Cumul des opportunités présentes dans les  $k$  zones les plus proches à  $i$

$O_N$  : Nombre total des opportunités

$p$  : Probabilité de choisir une opportunité dans une zone

Dans ce modèle, on pourrait considérer le nombre d'emplois disponibles en tant que le nombre d'opportunités pour le motif travail et le nombre de places scolaires comme le nombre d'opportunités pour le motif d'étude.

Pour montrer le défi relatif au modèle d'opportunités, on cite l'exemple suivant. Si on modifie l'offre de transport de façon à ce que l'ordonnancement des zones de destination à partir de chaque zone d'origine ne change pas, il n'y aura aucune incidence sur la matrice OD obtenue. Par contre, en utilisant le modèle gravitaire dans le même cas, il y aura des changements apportés aux nombres de déplacements estimés, car les temps de parcours sont explicitement présents dans la formalisation du modèle gravitaire. Normalement, cette incidence ne survient pas dans les cas réels avec un nombre de zones important.

Un autre enjeu relatif au modèle d'opportunités est sa sensibilité dans les cas où les intervalles de distance sont faibles. Cela peut arriver dans les zones centrales des villes, avec un nombre de zones et un niveau de service de transport élevés.

Dans les années 80, le modèle d'opportunités est appliqué à la région de Chicago (Illinois), pour réaliser les études de Chicago Area Transportation Studies, CATS (Eash, 1983, 1984).

Une approche récente pour modéliser la distribution des déplacements est l'approche dite consolidée, qui profite d'une formalisation mathématique d'optimisation multi objective. Cette approche intègre les contraintes représentant les modèles classiques tels que le modèle gravitaire (De Grange, Fernández, & De Cea, 2010).

Pour la calibration de la plupart des modèles de distribution, des matrices OD de référence sont requises. Plusieurs approches pour produire ces matrices sont présentées dans la littérature. Une approche coûteuse est de réaliser des enquêtes à bord des véhicules du TC. Une autre approche généralement moins coûteuse est d'estimer la matrice OD à partir de comptage des passagers montants et descendants aux arrêts et aussi des comptes à bord des véhicules (Blum, Sridhar, & Mathew, 2010; S. Feng & Li, 2004). Les grandes enquêtes OD des ménages représentent un autre moyen pour produire de telles matrices OD de référence (AMT, 2010) Une autre approche novatrice est d'utiliser les données provenant des cellulaires (Bonnell, Hombourger, Olteanu-Raimond, & Smoreda, 2014; Caceres, Wideberg, & Benitez, 2007; Herrera et al., 2010; Pan, Lu, Di, & Ran, 2006). Les données obtenues à partir de la technologie de Bluetooth (Blogg, Semler, Hingorani, & Troutbeck, 2010) et celles provenant des CAP sont également utilisées dans les développements récents (Munizagaa & Palmab, 2012). Une toute nouvelle étude novatrice a profité des données de localisation provenant des médias sociaux pour la détermination d'une matrice OD dans l'optique de développement d'un modèle gravitaire (Jin, Cebelak, Yang, Ran, & Walton, 2014).

La nécessité d'avoir une matrice OD de référence représente une limite majeure des modèles de distribution des déplacements, car dans beaucoup de cas, soit cette matrice n'existe pas, soit sa qualité n'est pas suffisamment bonne pour des fins de modélisation.

Après avoir calibré un modèle de distribution, il faut examiner sa validité d'ajustement et ce d'une façon systématique pour comparer la matrice estimée par le modèle avec celle de référence. De multiples mesures statistiques et approches spatiales sont présentées dans les Chapitres 5 et 6

de cette thèse. Ces chapitres contiennent également d'autres sujets d'intérêt relatifs à la distribution des déplacements en TC.

La troisième étape de la PSC est la répartition modale. Vu que cette thèse ne traite pas directement cette étape, seulement une brève revue est présentée.

### **2.1.1.3 Répartition modale**

Le but de cette étape est de segmenter les matrices OD produites au cours de l'étape précédente, par mode de déplacement. Même si l'intérêt de notre étude ne porte que sur les déplacements en TC, les principes relatifs au choix modal peuvent être intéressants au sein même des modes en TC. Autrement dit, de manière générale dans un réseau de TC, l'individu a la possibilité de choisir parmi un ensemble de modes tels que le train régional, le métro, le bus express, le bus régulier, le vélo, la marche, etc. Chaque mode de transport a des particularités qui représentent des utilités différentes pour des individus.

Les principales familles de modèles de répartition modale sont les modèles normatifs et les modèles Logit ou régression logistique (Bonnell, 2004). Les modèles Logit font partie d'une plus grande famille, les modèles de choix discrets. Le principe de ces modèles est que l'individu fait face à un univers de choix parmi lesquels il choisit l'alternative qui a l'utilité maximale pour lui (Train, 2009). L'enjeu essentiel relatif à l'étape de choix modal est la formalisation de la fonction d'utilité.

### **2.1.1.4 Affectation**

Cette étape consiste à affecter les déplacements OD estimés dans les étapes précédentes sur le réseau de transport, autrement dit le choix d'itinéraire. « *Le principe de base de tous les modèles d'affectation est de considérer que l'individu cherche à minimiser le coût généralisé de son déplacement pour une OD et un mode donnés* » (Bonnell, 2004). Différentes méthodes sont utilisées pour résoudre le problème d'affectation de déplacements, telles que l'affectation « *tout ou rien* », à l'équilibre (minimisation du temps par usager) et à l'équilibre de système (minimisation du temps total de déplacement). Des formalisations mathématiques accompagnées de modèles d'optimisation sont également utilisés à cet égard (Crainic & Florian, 2005; Sheffi, 1985). Nous citons ici certaines avancées plus particulièrement dans le domaine du TC.

La question d'affectation des déplacements dans sa forme générale est bien étudiée dans la littérature (Bonnell, 2004; Hensher & Button, 2008; Ortuzar & Willumsen, 1994). Par contre le

problème du choix d'itinéraires en TC est moins abordé. Hunt a présenté un modèle Logit pour estimer le choix d'itinéraires en TC. Il a utilisé une forme de coût généralisé qui représente l'influence de divers facteurs sur l'attractivité des alternatives dans un réseau de TC. Il a calibré le modèle en utilisant la procédure de maximum de vraisemblance décrite par McFadden. Dans son étude, l'accent est mis sur les déplacements motif travail à la destination du centre-ville à Edmonton (Alberta). Il a mentionné que son approche peut être généralisée pour divers motifs et localisations (Hunt, 1990; McFadden, 1974).

Les modèles de préférences déclarées (Stated Preferences ou SP) et préférences révélées (Revealed Preferences ou RP) sont également utilisés pour faire des études d'affectation des déplacements. En termes d'exemple, un modèle Logit binaire avec l'approche combinée de préférences déclarées et révélées est développé pour déterminer le choix d'itinéraires dans un réseau de bus à Yangon, Myanmar (Kato, Inagi, & Thet Htun, 2010). L'intérêt de ce modèle est l'utilisation combinée des données provenant des deux approches.

Iseki et al. ont étudié l'impact du temps en dehors de véhicules sur le comportement des voyageurs, particulièrement pour faire des correspondances. Ils ont développé des pénalités relatives aux correspondances afin d'évaluer l'influence des différents attributs de temps sur le comportement des voyageurs (Iseki, Taylor, & Miller, 2006).

Les pénalités relatives aux divers attributs de temps dans un réseau multimodal sont étudiées afin de déterminer le choix d'itinéraires dans un réseau de TC (Bovy & Hoogendoorn-Lanser, 2005). Également, Krygsman et al. ont examiné les attributs de temps de déplacement dans un contexte multimodal en TC. Ils ont conclu que l'importance des temps d'accès est plutôt associée à des modes d'accès en TC qu'à des caractéristiques sociodémographiques des individus (Krygsman, Dijst, & Arentze, 2004).

Une lacune majeure des approches d'affectation des déplacements en TC est qu'elles traitent la population en tant qu'un ensemble monolithique, ce qui ne permet pas de tenir compte des particularités, préférences et besoins spécifiques de différents segments de population. Par conséquent, il est d'un grand intérêt de formaliser cette question pour différentes catégories sociodémographiques, afin de valider leurs diversités comportementales. Nous reviendrons sur cette question dans le Chapitre 4.

L'approche totalement désagrégée présentée par MADITUC est également utilisée pour l'affectation des déplacements en TC (Chapleau, 1986). Nous présentons ces approches en détail dans la section 2.3.2.

Après avoir présenté la PSC dans sa forme générale, nous présentons la PSC intégrée sous deux volets distincts.

## **2.1.2 Procédure séquentielle classique intégrée**

Dans cette section, nous présentons deux catégories d'intégration de la PSC. La première consiste à intégrer certaines étapes de la PSC pour les réaliser d'une façon simultanée plutôt que séquentielle, et la deuxième vise l'intégration de la PSC avec d'autres familles de modèles notamment ceux démographiques et environnementaux pour le développement d'une plateforme universelle de la modélisation.

### **2.1.2.1 Intégration des étapes de la PSC**

Dans cette section, les études réalisées pour intégrer ou bien retirer certaines étapes de la PSC sont présentées. Dans une étude récente faite par Rajesson, un modèle hybride pour estimer la demande en TC est développé. Le modèle proposé a trois étapes principales : le traitement des données, la génération et la distribution des déplacements. La première étape consiste à rassembler et analyser les données sociodémographiques de la population et les caractéristiques du réseau de TC pour le territoire d'étude. La deuxième étape calibre les coefficients déterminant le nombre de déplacements entrants et sortants pour chaque zone par rapport à certaines variables indépendantes. L'estimation des matrices OD est réalisée dans la dernière étape (Rajesson, 2009). Une limite principale de cette approche est que l'auteur n'a utilisé que des variables indépendantes caractérisant le TC, alors que d'autres variables (motorisation, revenu, détention d'un permis de conduire, etc.) pourraient probablement être pertinentes. Cette approche montre une tentative pour la modélisation des déplacements effectués en TC, en se basant sur des comportements de l'utilisateur et en évitant l'étape de choix modal de la PSC.

Les modèles mathématiques sont aussi utilisés pour intégrer des étapes de la PSC. Galvan a développé une approche économétrique pour intégrer les étapes de distribution et d'affectation. Il a atteint une forme non linéaire, en intégrant le processus d'affectation dans l'estimation de la demande (Galvan, 1995). Une autre façon de calibrer un modèle de distribution-affectation intégré est l'utilisation d'algorithmes qui ne mettent de contraintes que sur les déplacements

produits. Cette approche de calibration n'est pas seulement plus rapide quant aux calculs, mais aussi nécessite moins de données de référence pour calibrer le modèle (Xu, Chen, & Gao, 2008). Zhou et al. ont développé une formalisation qui intègre les quatre étapes de la PSC, en utilisant des méthodes de programmation mathématique (Zhou, Chen, & Wong, 2009).

Une autre méthode pour développer un modèle de distribution-répartition modale intégré est l'utilisation des données de comptage des voyageurs. Un tel modèle est calibré avec les méthodes de moindre carrées et la maximisation de vraisemblance pour les données de Bandung, Indonésie (Tamin & Sulistyorini, 1997). L'avantage de ce modèle est sa simplicité d'application. Par contre, les résultats obtenus ne sont pas très satisfaisants, ce qui est probablement attribuable à la mauvaise qualité des données de référence.

Hossein Rashidi et Hasegawa ont intégré plusieurs étapes de la PSC dans l'optique du développement d'un modèle désagrégé (Hossein Rashidi & Hasegawa, 2014).

Ces études ont démontré quelques tentatives d'intégration des étapes de la PSC. L'intérêt d'intégration des étapes de la PSC est de diminuer l'effet séquentiel parfois négatif pour obtenir de meilleurs résultats. Dans la plupart des cas, les modèles mathématiques d'optimisation sont utilisés à des fins de calibration.

Dans la section suivante nous présentons quelques études réalisées pour intégrer des modèles démographiques et environnementaux avec la PSC.

### **2.1.2.2 Intégration de la PSC avec des modèles démographiques et environnementaux**

L'intérêt de cette intégration vient du fait que les changements démographiques dans le futur proche seront majeurs puisque selon Statistiques Canada, 20 % des canadiens auront plus de 65 ans en 2021 (StatCan, 2010). Jusqu'à présent, peu de recherches se sont appuyées sur les liens entre l'évolution de la structure démographique de la population et l'avenir des transports urbains. Une étude de Maoh et al. présente un système d'aide à la planification basé sur les Systèmes d'Information Géographiques (SIG) afin de voir les effets démographiques sur le transport urbain. Pour ce faire, un nouveau modèle nommé IMPACT (Integrated Model for Population Ageing Consequences on Transportation) est développé. L'objectif de ce modèle est d'intégrer les modèles démographiques et ceux de planification de transport pour pouvoir ensuite simuler certains scénarios démographiques et estimer leurs impacts sur les déplacements urbains. IMPACT est expérimenté sur la région de Hamilton (Ontario). La structure d'IMPACT a trois modules interconnectés qui sont les modules démographique, de transport et environnemental. Le

module démographique traite la population par sexe et par âge à l'aide de deux modèles distincts. Dans le module de transport, plusieurs profils sont étudiés avec la PSC : 8 profils pour la génération, 16 pour la distribution et 16 pour le choix modal des déplacements (par exemple 16 profils sont composés de 2 cohortes d'âge \* 2 motifs \* 4 périodes temporelles). Les 2 cohortes d'âge sont l'ensemble de la population qui conduit (âgée de 15 ans et plus) et les conducteurs adultes (âgée entre 15 et 65 ans). Les motifs sont travail et autre et les périodes temporelles sont le matin (6h00-9h00), la journée (9h00-16h00), l'après-midi (16h00-19h00) et le soir (19h00-6h00). Le dernier module estime les différentes émissions à partir des résultats obtenus du module précédent (Maoh, Kanaroglou, Scott, Paez, & Newbold, 2009).

IMPACT peut étudier les conséquences dues aux changements de :

- statistiques vitales comme les taux de natalité et mortalité ;
- caractéristiques de l'utilisation du sol comme la distribution d'emplois et l'accessibilité aux services ;
- infrastructures de transport comme la construction de nouvelles routes.

Également, IMPACT est capable de simuler les effets de changements socioéconomiques de la population sur le transport. Vu qu'IMPACT utilise des modèles statistiques pour la simulation, il y aura des erreurs dans le processus de prévision. Par contre, la comparaison des résultats obtenus de la simulation avec ceux réels pour 2001 et 2006 dans la région de Hamilton (Ontario) montre que le modèle se comporte bien. Les données requises pour l'utilisation de ce modèle sont celles des recensement et des enquêtes OD (Maoh et al., 2009).

Après avoir étudié les différents modèles en lien avec l'approche agrégée de la modélisation en transport, nous allons brièvement présenter l'approche désagrégée en TC.

## **2.2 Approche désagrégée**

Certains modèles et méthodologies présentés dans l'approche agrégée peuvent être traités d'une manière désagrégée. Nous présentons donc ici très brièvement les principes généraux relatifs à l'approche désagrégée et ses particularités principales par rapport à celle agrégée.



Dans l'approche désagrégée, les caractéristiques de chaque individu sont considérées d'une manière isolée. Cela distingue cette approche de celle agrégée qui considère un individu moyen et un système zonal.

Les modèles de choix discret ainsi que les modèles d'optimisation mathématiques sont couramment utilisés pour réaliser les développements en lien avec cette approche (Babazadeh, Khodakarami, & Zokaei Ashtiani, 2010; Vrtic et al., 2007).

L'approche désagrégée permet de mieux considérer des modèles dynamiques et des modèles qui tiennent compte des caractéristiques sociodémographiques. Également, en appliquant la démarche désagrégée, l'analyste est en mesure de définir le niveau d'agrégation a posteriori selon l'objet d'étude (Hutchinson, 1974; Ortuzar & Willumsen, 1994).

## **2.3 Approche descriptive**

L'approche descriptive est une approche de planification qui repose sur l'utilisation et l'analyse des résultats obtenus des enquêtes en transport. Plus généralement, cette approche est fondée sur la recherche de types de comportements, d'indicateurs moyens et de corrélations entre certains facteurs représentant la demande en transport (Cirillo, Cornélis, Hubert, & Toint 2004; ECMT, 1999).

Dans cette revue, nous classons tous les modèles qui sont essentiellement basés sur les données de transport dans la famille des modèles descriptifs. Le point commun entre ces modèles est l'initialisation par les observations provenant de différentes sources et le suivi avec le calcul des paramètres descriptifs statistiques pour décrire les caractéristiques de la mobilité. L'objectif général est de reproduire le mieux possible les observations par le biais de modèles descriptifs.

Étant donné le rôle essentiel des données dans cette approche, une bonne connaissance des sources de données en TC est de grand intérêt. Cette connaissance permet également l'utilisation de ces données dans d'autres approches de modélisation en TC. Dans le Chapitre 8, nous discuterons de la possibilité d'arrimage des méthodologies de modélisation et des nouvelles données en tant que piste de réflexion pour le futur.

Cette section consiste à discuter de différentes sources de données ainsi que des applications et des développements méthodologiques en lien avec elles.

### 2.3.1 Sources de données

De multiples sources de données sont exploitées dans les diverses études en TC. En voici quelques-unes :

- Recensement : Ces données contiennent divers attributs sur la population dont le nombre, l'âge, le sexe, la taille et la composition des ménages, le revenu, la langue, etc. Les découpages territoriaux du recensement canadien sont fins allant jusqu'à des unités d'environ 1 000 personnes.
- Inventaires d'emploi et de places scolaires : Il s'agit du nombre d'emplois et de places scolaires, soit par zone, soit d'une façon désagrégée.
- Enquête OD : Les enquêtes OD sont les méthodes « *qui permettent l'obtention d'un portail global des habitudes de déplacements d'une population dans une journée moyenne* » (Trépanier, Morency, & Blanchette, 2008). Les données recueillies touchent les attributs des ménages (taille, motorisation, lieu de domicile, etc.), des personnes (âge, sexe, permis de conduire, etc.) et des déplacements (origine, destination, heure, séquence de modes, motif, etc.).
- Enquête à bord : Ces enquêtes effectuées à bord des véhicules de TC peuvent avoir des objectifs divers, ce qui se traduit par des formes différentes d'enquête. Elles peuvent avoir la forme générale d'enquêtes OD ou bien d'autres formes, par exemple pour apprécier le niveau de satisfaction de la clientèle, ou pour étudier l'impact d'une modification de l'offre sur le comportement des usagers en TC.
- Caractéristiques de l'offre du TC : Plusieurs types de données sont reconnus. Les données de General Transit Feed Specifications (GTFS) représentent une source complète pour la caractérisation de l'offre planifiée du TC.
- Localisation des véhicules de TC : Les données provenant des systèmes d'Automatic Vehicle Location (AVL) et de Global Positioning System (GPS) peuvent être utilisées comme des compléments pour d'autres sources déjà présentées. Ces outils permettent d'obtenir des informations continues, détaillées et précises sur la localisation de véhicules de TC.
- Compte à bord : Le comptage des passagers à bord des véhicules permet de déterminer la charge sur les lignes de TC à des périodes et endroits donnés. Cette opération peut être

réalisée par les observateurs ou bien par des systèmes automatiques de comptage des passagers tels qu'un Automatic Passenger Counting (APC). La fusion de ces données avec celles de localisation des véhicules permet d'avoir un portrait plus précis sur la charge à bord des véhicules dans le temps et l'espace, spécifiquement dans le cas de collectes de données automatiques.

- Systèmes de paiement par cartes à puce (CAP) : Depuis des années, les CAP sont utilisées par les fournisseurs du TC comme des moyens pour effectuer des paiements. Ce système de paiement permet de recueillir les données concernant les déplacements effectués dans un contexte urbain. Ces données archivées pourront donc être utiles pour des fins d'analyse descriptive et de modélisation en TC (Pelletier, Trépanier, & Morency, 2009).
- Données territoriales : Les données territoriales sont les données qui décrivent le territoire d'étude. Les Systèmes d'Information Géographiques (SIG) sont utilisés pour la manipulation de ces données ainsi que l'intégration d'autres données avec celles-ci.

Après avoir identifié les diverses sources de données, nous étudions MADITUC et l'approche orientée-objet suivi par la présentation de certaines applications en lien avec d'autres données en TC.

### **2.3.2 Enquête OD, MADITUC et l'approche orientée-objet**

L'approche totalement désagrégée en transport urbain présentée par MADITUC permet de manipuler la totalité de l'information déclarée au cours d'une enquête OD (Chapleau, 1986). Étant donnée la diversité importante des attributs collectés lors d'une enquête, l'utilisation d'une approche orientée-objet pour la manipulation des données et la réalisation des études est bénéfique. Cette approche qui permet également d'identifier de nouveaux objets sert à mieux exploiter les données d'enquêtes (Trépanier & Chapleau, 2001). Vu le rôle principal des enquêtes OD en lien avec MADITUC et l'approche orientée-objet, nous expliquons tout d'abord brièvement quelques constats relatifs aux enquêtes OD.

Les enquêtes ménages OD sont des outils qui produisent des sources de données importantes en matière de mobilité des personnes (CIMTU, 2010). Les premières tentatives pour réaliser les grandes enquêtes OD ont été faites dans les années 50 aux États-Unis (Cirillo et al., 2004). Cette pratique est suivie plus tard avec les enquêtes OD régionales canadiennes à Montréal (Québec) et

à Toronto (Ontario), l'enquête nationale des Pays-Bas, l'enquête de Transport et Communications en France, la Nation-wide Personal Transportation Survey (NPTS) des États-Unis, le micro recensement suisse, l'enquête KONTIV en Allemagne, l'enquête nationale belge Mobel, et la grande enquête OD de Téhéran (Iran) (Bonnel, 2004; Cirillo et al., 2004; Zakeri Sohi, 2007). Nous rappelons que ces enquêtes permettent généralement de recueillir des informations sur trois niveaux : les ménages, les personnes et les déplacements (Richardson, Ampt, & Meyburg, 1995). Récemment quelques enquêtes différentes sont réalisées par internet, nommées les enquêtes web (Morency, Verreault, & Bourbonnais, 2013; Viggiano, Koutsopoulos, & Attanucci, 2014; Volosin et al., 2014).

La réalisation des analyses au niveau des ménages permet de connaître certains aspects plus généraux comme l'état de motorisation et d'accessibilité aux différents modes de transport à partir du lieu de résidence. La collecte des données relatives aux personnes clarifie la mobilité des individus comme les chaînes d'activités et de déplacements. Le troisième niveau qui englobe les déplacements, nous informe sur les trajets effectués entre deux lieux, à une heure donnée, par un ou des modes de transport déclarés, et ce pour un motif précis (Cirillo et al., 2004).

Dans la grande région de Montréal (Québec) les enquêtes ménages OD font partie de la culture de la planification en transport depuis 1970. Elles permettent de « *recueillir une information détaillée sur tous les déplacements effectués par les individus d'un ménage, sur un échantillon d'environ 5 % des ménages de la région montréalaise* » (Chapleau, Allard, Trépanier, & Morency, 2001).

À l'aide de MADITUC et de l'approche orientée-objet, au lieu de mettre les informations dans une matrice OD qui « *agrègerait et dissoudrait l'information* », les données des ménages, personnes et déplacements sont analysées d'une façon désagrégée en laissant la possibilité d'agrégation au besoin (Trépanier, Chapleau, & Morency, 2008). La possibilité d'agrégation a posteriori « *permet de proposer un découpage du territoire adapté à l'objet d'étude* » (Thévenin, 2001).

Le groupe MADITUC de l'École Polytechnique de Montréal a développé des logiciels pour recueillir et traiter les données nécessaires à la planification des transports. À l'aide de l'approche MADITUC et les données provenant des enquêtes OD, quatre types d'applications sont envisageables : le développement des modèles d'analyse du réseau, la réalisation des analyses multi perspectives sur la demande en transport, l'évaluation et l'estimation des effets

géopolitiques en transport ainsi que le développement des modèles orientés-objets et des applications de dissémination d'informations pour les usagers (Chapleau, 2002).

L'approche orientée-objet est de plus en plus utilisée dans divers domaines scientifiques en raison des similarités qui existent entre cette approche et la réalité des phénomènes; elle présente aussi l'avantage de la réutilisabilité des composants d'un problème (Trépanier & Chapleau, 2001). Dans le domaine de transport, quatre classes d'objets sont identifiées (Trépanier, 1999) :

- Statique : Ces objets ont une localisation fixe dans le temps et l'espace (zones, codes postaux, secteurs de recensement, etc.).
- Dynamique : Cette classe d'objets englobe les acteurs de transport. Autrement dit, ces objets décident et contribuent à leur mouvement (individus, modes de transport comme bus et automobile, marchandises, etc.).
- Cinétique : Ces objets sont les descripteurs du mouvement (déplacement, trajet, ligne de TC, etc.).
- Systémique : Cette classe représente un groupe d'objets inter-reliés qu'ils soient opérationnels (réseau de TC, réseau routier, etc.), informationnels (enquête, recensement, etc.) ou multifonctionnels (centre-ville, banlieue, etc.).

Dans le processus d'examen d'une enquête OD, Trépanier et al. ont déterminé le modèle objet enrichi. Ce modèle est dit enrichi car il présente certains objets dérivés ou créés comme le statut des personnes et des générateurs de déplacements à partir des données brutes de l'enquête (Trépanier & Chapleau, 2001).

L'approche orientée-objet est reconnue comme l'extension de MADITUC dans la littérature (Chapleau & Trépanier, 1997). Dans une étude effectuée par Trépanier et al. sur les générateurs de déplacements, le Modèle d'Analyse Désagrégée des Générateurs de Déplacements (MADGEN) est présenté et examiné avec les données des enquêtes OD 1987 et 1993 de Montréal (Québec) (Trépanier & Chapleau, 1996). D'autres études ont traité les différentes caractéristiques relatives aux lieux d'emplois et aux travailleurs (Chapleau & Lavigueur, 1996). Les effets redistributifs et les problématiques fiscales de l'usage de TC sont également discutés grâce à ces approches (Dalil Essakali & Chapleau, 2000).

L'intégration de MADITUC et de l'approche orientée-objet avec une approche floue (Fuzzy) permet d'éviter quelques problèmes dus au manque de données. Une étude de Chapleau et al.

présente le passage graduel des niveaux d'information d'un niveau spatial haut à un niveau spatial fin (Chapleau, Trépanier, & Allard, 1998).

L'étude de MADITUC et de l'approche orientée-objet démontre le potentiel des données provenant des enquêtes OD pour des fins d'analyses et de modélisations descriptives. Dans la section suivante, nous présentons le potentiel d'autres types de données en TC.

### **2.3.3 Enquêtes à bord des véhicules**

Les enquêtes à bord des véhicules de TC sont effectuées depuis longtemps pour des buts divers. Le processus de collecte de données a été amélioré au cours des années. Une étude à Madison (Wisconsin) a montré que certains éléments comme la durée de l'enquête peuvent affecter la qualité des résultats (Komanduri & Proussaloglou, 2010). L'utilisation des nouvelles technologies comme des assistants numériques personnels (Personal Digital Assistant, PDA) équipés de GPS peut également améliorer la qualité des informations recueillies. Une expérimentation à Columbus (Ohio) a démontré l'efficacité de l'utilisation des nouvelles technologies dans le processus d'une enquête à bord (Simas Oliveira & Casas, 2010). Ces enquêtes sont utilisées de façons différentes pour réaliser les analyses descriptives et la modélisation en TC.

En 1951, la Commission de Transport de Montréal (CTM) a réalisé une grande enquête à bord de ses autobus et tramways à Montréal (Québec). En 1961, ils ont effectué une autre enquête auprès des usagers de TC qui a servi plus tard à l'établissement des premières lignes de métro montréalaises (Morency & Trépanier, 1998).

L'analyse de la fiabilité d'un réseau de bus est effectuée à l'aide des données provenant d'une enquête à bord à Beijing (Chine) (X. Chen, Yu, Zhang, & Guo, 2010). Une lacune majeure de cette étude est la courte durée de l'enquête qui était seulement de trois jours.

De multiples analyses comportementales sont également possibles à partir des informations provenant de ce type d'enquête. Currie et Loader ont examinés l'impact des correspondances sur le comportement des voyageurs du réseau de bus de Melbourne (Australie) à l'aide d'une enquête à bord (Currie & Loader, 2010).

La réalisation des études comportementales notamment le comportement de choix est également un sujet d'intérêt majeur avec les données provenant des enquêtes à bord. Cela s'inscrit dans des études de modélisation de choix à partir des préférences déclarées des voyageurs en TC.

Généralement la validation des données de ces enquêtes est un processus long. Dans la littérature récente un outil source ouverte développé sous ArcMap d'ESRI est présenté pour la validation d'une enquête à bord à partir des données de localisation des véhicules. Cette approche de correction des cas improbables a permis d'accélérer le processus de validation (Rohne, 2014).

La section suivante présente le potentiel des données opérationnelles et passives dans l'analyse descriptive et la modélisation en TC.

### **2.3.4 Potentiel des données opérationnelles et passives**

Cette section aborde une revue de littérature sur l'utilisation des données opérationnelles et passives telles que les données de GTFS, AVL, GPS, compte à bord et CAP.

La caractérisation des déplacements permet une meilleure compréhension des comportements des voyageurs. Les méthodes plus couramment utilisées comme la dérivation de comportements à partir des enquêtes OD comportent plusieurs défis relatifs à la pondération et au contrôle des biais des enquêtes (Riegel & Attanucci, 2014; Spurr, Chapleau, & Piché, 2014; Toronto, 2008). Les données du service planifié peuvent être utilisées pour déterminer plusieurs éléments en planification du TC, dont le niveau de service du TC planifié. Les données opérationnelles et passives peuvent contenir également des informations précises et détaillées sur les comportements de mobilité. Par contre, ces dernières ont des limites relatives à la qualité des données ou l'absence de certains attributs des déplacements (Chu & Chapleau, 2010).

Cette section vise à présenter brièvement trois catégories de 1) données de services planifiés, 2) validations par CAP et 3) systèmes de comptage et localisation des véhicules de TC.

#### **2.3.4.1 Données de services planifiés**

Les données relatives au service planifié de TC peuvent avoir des formes différentes. Dans cette thèse, les données GTFS sont utilisées pour représenter le niveau de service planifié du TC (détails présentés aux Chapitres 5 et 6).

Le format de GTFS a été développé par Google et TriMet, l'Autorité Organisatrice de Transport (AOT) de l'agglomération de Portland (Oregon) en 2005. Cela représente un format informatique standardisé pour intégrer les horaires et les informations géographiques relatifs à un réseau de TC (Google, 2012; Wikipédia, 2014).

Cette plateforme permet aux AOT à travers le monde de communiquer leur service de TC planifié à Google pour l'intégrer dans le calculateur de trajet de Google Maps. Le régime permanent de production de ces données généralement disponibles pour le grand public ainsi que leur format standardisé unique présentent une opportunité pour les intégrer aux analyses descriptives et de modélisation en TC. Certains chercheurs ont récemment bénéficié de ces données dans divers domaines de la planification en TC (Gandavarapu, 2012; Godin, 2012; Martel Poliquin, 2012; Wong, 2013).

#### **2.3.4.2 Données passives de validations par CAP**

Les CAP sont similaires aux cartes de crédits du point de vue de la forme et de la taille ; elles peuvent remplacer les systèmes de paiements traditionnels tels que les billets en papier dans un réseau de TC. Ces cartes peuvent être personnelles (non transférables à un autre usager) comme à Gatineau (Québec), ou anonymes comme à Montréal (Québec) et utilisées comme un titre de transport ou comme une carte pour y charger de l'argent et en débiter au moment du déplacement. Les données de CAP ont été utilisées pour des fins d'analyses descriptives et de modélisation en TC (Bagchi & White, 2004; Chu, 2014; Chu & Chapleau, 2010; Jang, 2010; Morency, Trépanier, & Agard, 2006a, 2007; Park & Kim, 2008; Shi & Lin, 2014; Tranchant, 2005; Uniman, Attanucci, Mishalani, & Wilson, 2010; Utsunomiya, Attanucci, & Wilson, 2006). Voici quelques avantages des données de CAP (Bagchi & White, 2004) :

- accès à des données plus détaillées concernant les déplacements faits sur le réseau ;
- possibilité d'associer les déplacements aux cartes individuelles afin de pouvoir suivre le comportement de mobilité des cartes, donc de personnes mais non identifiables, dans le temps et l'espace.

Rappelons que généralement deux types de transactions sont faites avec les CAP. Le premier survient au moment où le détenteur de la carte fait un déplacement. À ce moment-là, certaines informations concernant le déplacement effectué dont le lieu d'embarquement, l'heure et la date sont enregistrées. Le deuxième type est la transaction monétaire au moment du chargement de la carte ou bien de l'achat d'un titre de transport.

Dans la plupart des systèmes de paiement par CAP, les voyageurs ne valident leur carte qu'aux points d'embarquement ; il s'agit d'une simple validation. Dans certains systèmes, les voyageurs



valident la carte à la fois aux points d'embarquement et de débarquement. Il y a donc deux validations qui peuvent produire des informations très détaillées et complètes quant aux origines et destinations des déplacements sur le réseau, (Park & Kim, 2008). Par contre cette approche est moins acceptée par les voyageurs, probablement à cause de diminuer la rapidité et le confort au moment de débarquement. Une étude en Angleterre montre que dans un système à double-validations sur une période d'essai, un tiers des usagers n'ont pas validé leur carte aux points de débarquement (Merseytravel, 2001). Afin de profiter des données provenant des systèmes de simple-validation comme dans le cas de Montréal (Québec), l'analyste aura besoin d'enrichir les données collectées en estimant les points de débarquement.

Trépanier et al. ont développé une méthodologie de dérivation des destinations qui a été testée pour le réseau de Gatineau (Québec) avec un taux de réussite d'environ 80 % pour estimer les destinations durant les heures de pointe (Trépanier, Tranchant, & Chapleau, 2007).

Chu et Chapleau ont également réalisé des travaux de recherche sur l'enrichissement des données de CAP. Ils ont examiné les données de la Société de Transport de Laval (STL) sur une période d'un mois, dont 713 276 validations dans 240 bus. Pour la caractérisation des déplacements, les points d'ancrage qui sont les points souvent visités par chaque carte sont identifiés. Leur étude a démontré que malgré le manque de certains attributs comme le motif de déplacement, l'enrichissement des données permet de comprendre le comportement de mobilité des voyageurs (Chu & Chapleau, 2010). L'enjeu majeur concerne la possibilité d'automatisation de tels processus d'enrichissement.

D'autres études pour enrichir des données de CAP en lien avec les méthodologies de « *data mining* » (fouille systématique de données) sont présentées dans la littérature. Ces études permettent d'avoir une meilleure compréhension des rythmes d'activités en analysant les données de CAP (Kieu, Bhaskar, & Chung, 2014; Morency et al., 2006a; Morency, Trépanier, & Agard, 2006b; Morency et al., 2007).

Un autre enjeu relatif à l'utilisation des données de CAP est que certains voyageurs n'utilisent pas de CAP. Le développement de certaines méthodologies pour corriger le biais relatif à ces voyageurs semble être bénéfique.

Globalement deux types d'analyse sont réalisables à l'aide des données de CAP : la caractérisation de la demande et la caractérisation de l'offre de TC. Le premier vise à étudier le comportement de mobilité des voyageurs et le deuxième consiste à analyser entre autres la

performance et le niveau de service en TC. Nous discutons de quelques études présentées dans la littérature.

L'appréciation du comportement des usagers est un élément indispensable quant à la planification du TC. Les données passives présentent une piste exploratoire intéressante à cette fin. Plusieurs chercheurs et professionnels ont utilisé ces données de façons différentes.

Depuis 2004, la ville de Séoul (Corée du Sud) profite d'un système tarifaire basé sur la distance parcourue qui nécessite la double-validation des CAP aux points d'embarquement et de débarquement. Jang a montré le potentiel de ces données pour le calcul et la comparaison de temps de déplacement entre les arrêts en différents modes et également l'analyse des correspondances (Jang, 2010).

Dans une autre étude, Guo et Wilson ont étudié l'impact des correspondances sur le choix d'itinéraires dans le métro de Londres (Angleterre). Ils ont découvert que la prise en considération d'une pénalité globale pour toutes les correspondances différentes ne peut pas bien représenter le comportement des voyageurs dans un réseau de TC. Ils ont validé également que les voyageurs préfèrent les trajets plus longs avec moins de correspondances aux trajets plus rapides avec un nombre de correspondances plus élevé (Guo & Wilson, 2010).

L'analyse de la performance d'un système de TC est la première étape vers une gestion efficace et proactive (Bertini & El-Geneidy, 2003). L'utilisation des données passives diminue la nécessité de faire des hypothèses sur l'uniformité ou la généralisation du comportement des voyageurs (Berkow, El-Geneidy, Bertini, & Crout, 2009). Utsunomiya et al. ont présenté certains potentiels d'utilisation des données de CAP pour des fins d'amélioration de la planification en TC (Utsunomiya et al., 2006). Trépanier et al. ont calculé certains indicateurs de performance à la fois pour l'offre (véhicule-kilomètre, véhicule-heure, vitesse commerciale, etc.) et la demande (passager-kilomètre, passager-heure, distance et durée moyennes de déplacement, etc.) de TC, à partir des données de Gatineau (Québec) (Trépanier, Morency, & Agard, 2009). Cette approche permet de faire des analyses très fines aux niveaux temporels et spatiaux, mais elle est limitée notamment au niveau technique, vu la quantité de données à traiter.

Un autre sujet d'intérêt majeur est l'étude de la fiabilité d'un service de TC à l'aide des données opérationnelles et passives. Dans une étude réalisée par Uniman et al., les mesures de fiabilité sont examinées à l'aide des données de CAP (Uniman et al., 2010).

Ces études démontrent le potentiel des données de CAP pour des fins d'analyse descriptive et de modélisation en TC. Trépanier et al. supposent que les données passives comprenant des niveaux de résolution temporels et spatiaux fins sont complémentaires aux données d'enquêtes OD qui offrent une vision détaillée de tous les déplacements faits par un échantillon de la population (Trépanier, Morency, & Blanchette, 2008).

### **2.3.4.3 Données de systèmes de comptage et de localisation des véhicules**

Les données des systèmes de comptage des voyageurs manuels et automatiques ainsi que celles de localisation des véhicules peuvent être utilisées dans l'analyse et la modélisation descriptive du TC. Dans cette section, certaines études présentées dans la littérature sont expliquées.

Les données provenant d'un système d'APC crée des informations continues sur les embarquements et les débarquements. L'avantage de ce système automatique par rapport aux systèmes manuels est qu'il permet la collecte des données sur les grands réseaux de manière continue dans le temps et l'espace. McCord et al. ont développé une méthodologie pour la dérivation d'une matrice OD à partir des données d'APC sur une ligne de bus. L'évaluation des résultats obtenus à partir de cette méthodologie par rapport à ceux dérivés d'une enquête à bord a confirmé que l'intégration des deux sources de données produit de meilleurs résultats (McCord, Mishalani, Goel, & Strohl, 2010).

L'utilisation simultanée des données provenant d'un système de compte à bord et d'une enquête OD pour la détermination des matrices OD est également étudiée à Mumbai (Inde). Le processus développé est nommé « *l'estimation de la demande hybride* » (Blum et al., 2010).

Les données de localisation des véhicules présentent une autre source de données bénéfiques. Cortés et al. ont utilisé un ensemble de données concernant la localisation automatique des bus afin de déterminer la vitesse commerciale sur un réseau en convertissant un volume de données important en des informations utilisables (Cortés, Gibson, Gschwender, Munizaga, & Zuniga, 2010).

L'étude de la ponctualité de services de TC à l'aide des données d'AVL et d'APC est de grand intérêt pour les fournisseurs de TC (Furth, Hemily, Muller, & Strathman, 2006). Mandelzys et Hellinga ont développé une méthode pour l'identification automatique des causes de non-respect des horaires planifiés dans un réseau de bus. À l'aide des données d'AVL et d'APC, ils ont identifié des catégories de problème et trouvé des arrêts problématiques

(Mandelzys & Hellinga, 2010). Le point fort de cette étude est l'automatisation de la méthodologie. Par contre, les volumes de passagers ne sont pas associés aux arrêts et donc l'analyse de l'importance des problèmes en fonction des volumes de passagers n'est pas faisable. Plusieurs autres études sont réalisées en lien avec l'enrichissement des données obtenues des systèmes de GPS en ajoutant certains attributs de déplacements tels que le motif (Arifn & Axhausen, 2012; Montini, Rieser-Schüssler, Horni, & Axhausen, 2014).

La section suivante présente globalement les systèmes d'aide à la planification.

### 2.3.5 Systèmes d'aide à la planification

Les Systèmes d'Information Géographiques (SIG), tel un système d'aide à la planification, «*facilitent l'intégration des données sur le territoire, les réseaux routiers et les réseaux de transport avec les techniques de modélisation* » (Chapleau & Morency, 2005).

Les SIG sont utilisés dans la plupart des études en planification de TC pour intégrer les données provenant d'autres sources dans un contexte géographique (C. Chen, Gong, Lawson, Bialostozky, & Muckell, 2010). Ils peuvent être également utilisés en tant qu'outil primaire pour faire des «*sketchs plans* » en lien avec les problèmes de design de réseau de TC. Une expérimentation à Auburn (Alabama) montre le potentiel des SIG pour améliorer l'efficacité d'un système de bus scolaire. L'avantage de cette approche est la simplicité d'application (Chaudhari & Ye, 2010).

Simard et al. ont examiné l'utilisation des SIG pour le design et l'analyse des réseaux de TC à l'aide de nouveaux outils informatiques développés sous Visual Basic for Applications (VBA) dans le logiciel ArcMap d'ESRI. L'objectif est d'analyser les corridors de TC pour contribuer à résoudre le problème de design de réseau. La méthodologie est basée sur l'identification des attributs de l'utilisation du sol (nombre d'emplois, etc.) et de la démographie (population, etc.) au niveau des zones d'étude à l'aide des SIG. Par la suite, l'outil permet à l'analyste de fusionner des zones pour créer des pôles d'activités au sein desquels les nombres de déplacements sont plus élevés. Après avoir fait cette agrégation, l'outil fournit une représentation visuelle pour montrer les volumes de voyageurs. À l'aide de cette visualisation, l'analyste peut formuler différents alternatives pour le nouveau réseau de TC en lien avec la population à desservir. La dernière étape consiste à calculer des coûts opérationnels associés à chaque alternative afin de choisir la meilleure. Cette approche est expérimentée dans la région de Waterloo (Ontario) (Simard, Springate, & Casello, 2010).

Au cours de ces dernières années, grâce aux avancées technologiques et la vaste utilisation de télécommunications comme les téléphones intelligents (smartphone), d'autres sources de données sont reconnues. Parmi les études récentes on peut citer le calcul de temps de parcours à l'aide des données provenant d'un système de Bluetooth (Malinovskiy, Wu, Wang, & Lee, 2010) et le développement des matrices OD à partir des données de cellulaires (Bonnell et al., 2014; Pan et al., 2006). Étant donné la quantité très élevée de certaines nouvelles données, les approches d'analyse et de traitement de « *big data* » sont également discutées dans la littérature récente (IBM, 2013; OECD, 2013). Par la définition, « *big data* » est un terme générique pour un ensemble de données très complexe et difficilement traitable. Les défis comprennent entre autres le stockage, la recherche, le partage, le transfert, l'analyse et la visualisation (Wikipédia, 2014). L'étude de différentes sources de données en lien avec les approches de modélisation notamment celle descriptive permet également d'identifier leur potentiel d'utilisation dans le processus de la modélisation en TC.

## **2.4 Approche de simulation**

Dans cette section, nous présentons deux types de simulation dans les études de transport : la simulation des données de référence et l'utilisation de la simulation pour évaluer les scénarios.

### **2.4.1 Simulation des données de référence**

La simulation des données de référence consiste généralement à créer une population fictive distribuée sur le territoire d'étude. Certaines tentatives pour développer les modèles de simulation de ménages et de personnes sont présentées dans la littérature dont 1) le système Household Attributes Generation System (HAGS) du modèle Florida Activity Mobility Simulator (FAMOS) en Floride et 2) le modèle développé à l'université d'Illinois à Chicago (Javanmardi, Rashidi, & Mohammadian, 2010; Pendyala, Kitamura, & Kikuchi, 2004).

FAMOS est un modèle de simulation qui prend en compte le niveau individuel dans le processus de prise de décision. Ce modèle génère des ménages et des individus avec leurs attributs en utilisant des données sociodémographiques et socioéconomiques et celles provenant des enquêtes OD. FAMOS simule également les activités principales telles que les déplacements motif travail et étude (Pendyala et al., 2004). La lacune majeure de FAMOS est qu'il ne prend en

considération que les déplacements de ces deux motifs. Javanmardi et al. ont développé un autre simulateur de population et des attributs de déplacements, et ils l'ont examiné dans les cas de Des Moines (Iowa) et l'état de New York (Javanmardi et al., 2010). Ces nouveaux simulateurs de population font le transfert de données et de caractéristiques comportementales de mobilité d'une population réelle vers une population fictive.

Nous supposons que l'approche de simulation des données de références est de plus en plus intéressante, ce qui est attribuable à la fois à la complexité et aux coûts importants associés à la réalisation des enquêtes OD. De plus cette approche permet d'étudier certains scénarios tels que le vieillissement de la population et la croissance de la présence des femmes sur le marché de travail.

## **2.4.2 Simulation des scénarios en transport**

La simulation des scénarios est de plus en plus utilisée dans la pratique par les planificateurs de TC. Dans cette section, certaines études de simulation sont brièvement présentées afin d'identifier le potentiel et les limites de cette approche.

La simulation peut servir aux études de correspondances dans un réseau de TC. Elle représente un outil performant pour déterminer l'efficacité des politiques de synchronisation et de planification dynamique des correspondances (Hadas & Ceder, 2008).

Une autre application est de simuler des correspondances avec des flux de voyageurs pour déterminer les problématiques associées. Wang et al. ont étudié des correspondances dans un terminal à Beijing (Chine). Les résultats de simulation des flux de voyageurs avec le logiciel VISSIM a démontré que la diminution de la distance de marche et l'amélioration des infrastructures piétonnières sont des éléments essentiels pour encourager les usagers à faire des correspondances (Wang, Chen, He, & Gao, 2006).

Bachelet et Yon ont développé une méthode hybride pour le design d'un réseau de bus et les horaires associés. Compte tenu de la complexité du problème, l'utilisation d'une méthode d'optimisation mathématique sur l'ensemble du réseau semble être lente. Face à ce problème, ils ont développé un modèle amélioré qui intègre l'approche mathématique et celle de simulation. Pour ce faire, le modèle mathématique est appliqué sur une partie simplifiée du réseau et ensuite les résultats sont généralisés à l'aide d'une simulation. Le modèle hybride permet d'obtenir une

bonne solution avec un temps de calcul raisonnable. La lacune majeure de cette approche hybride est que la convergence n'est pas toujours garantie (Bachelet & Yon, 2005).

La simulation est utilisée pour évaluer l'impact des stratégies de priorisation de passage des véhicules de TC dans les intersections ainsi que les corridors exclusifs pour les bus à haut niveau de service (Morgan, 2002; Siddique & Khan, 2006). Ces études montrent la puissance de l'approche de simulation dans divers domaines d'étude en planification du TC. Cependant, la simulation ne sert habituellement qu'à examiner des problématiques précises et locales.

## 2.5 Approche opérationnelle

La Planification Opérationnelle en TC (POTC) est l'étape en aval de la prévision de la demande. Autrement dit, pour réaliser la POTC on aura besoin des résultats des modèles de prévisions de la demande. Bien que la POTC ne soit pas étudiée dans cette recherche, elle est brièvement présentée dans la revue de littérature afin d'avoir un regard plus global sur la problématique de planification en TC dans ses différents aspects. Cette section est basée principalement sur la démarche présentée dans l'ouvrage « Public Transit Planning and Operation: Theory, Modeling and Practice » (Ceder, 2007).

La POTC consiste à réaliser quatre étapes principales (Ceder, 2007) :

- design de réseau composé des lignes, arrêts, terminaux, etc. ;
- développement des intervalles de service ;
- développement des horaires ;
- affectation du personnel comme l'assignation des chauffeurs, etc.

Ces étapes peuvent être catégorisées selon les niveaux de planification dont stratégique, tactique et opérationnel démontrés à la Figure 2.5 (Ceder, 2007; Guihaire & Hao, 2008; Pelletier et al., 2009).

Les principaux objectifs du design de réseau de TC sont d'avoir une couverture nécessaire du territoire, de maximiser les itinéraires directs et de satisfaire la demande. La deuxième étape consiste à déterminer la fréquence de départ sur des lignes de TC. Dans la troisième étape, on s'assure d'avoir une coordination, dans la mesure du possible, entre les horaires de différentes lignes dans chaque correspondance pour améliorer le niveau de service global. La dernière étape

est consacré à l'affectation des chauffeurs sur une journée et aussi à développer l'affectation du personnel sur une période plus longue.

Guihaire et Hao ont effectué une revue globale sur le problème de design de réseau et de développement des horaires des véhicules de TC. Leur revue porte sur les niveaux stratégique et tactique d'étude en POTC. Dans leur revue, différentes approches mathématiques et heuristiques relatives à chacune de ces deux étapes sont présentées. Certaines études présentées dans cette revue consistent à résoudre les deux étapes d'une manière simultanée. Généralement, les approches mathématiques sont difficilement applicables aux grands réseaux réels en raison de la complexité des calculs; ceci explique l'intérêt d'utiliser des approches heuristiques (Guihaire & Hao, 2008).

La réalisation de chaque étape de la POTC nécessite l'acquisition de certaines données. Ceder a identifié cinq méthodologies différentes de collecte de données, dont le contrôle des nœuds de réseau, le contrôle à bord des véhicules, le contrôle de kilométrage et de temps mort, l'enquête de voyageurs et l'enquête de population (Ceder, 2007). Plusieurs études ont démontré l'utilisation des données de types divers dans le processus de la POTC (Cevallos, Wang, Gan, & Chen, 2010; Dessouky, Hall, Nowroozi, & Mourikas, 1999; Fu, Liu, & Calamai, 2003).

Après avoir présenté les cinq approches de la modélisation en transport, nous faisons une synthèse de la revue de littérature, suivie par une discussion en liens avec les limites constatées. Ceci permet de clarifier l'intérêt de la réalisation de cette recherche pour contribuer à la modélisation des déplacements en TC.

## **2.6 Synthèse et discussion**

Au cours de cette revue de littérature, les différentes approches de la modélisation en TC ont été identifiées et discutées. Étant donné le nombre importants des références associées, le Tableau A.1, en Annexe A, présente sommairement une synthèse chronologique des chercheurs et de leurs études concernant les différents aspects de la modélisation en TC.

Le Tableau 2.1 résume les points forts autrement dit le potentiel d'utilisation, et les points faibles ou encore les limites des approches présentées dans cette revue.



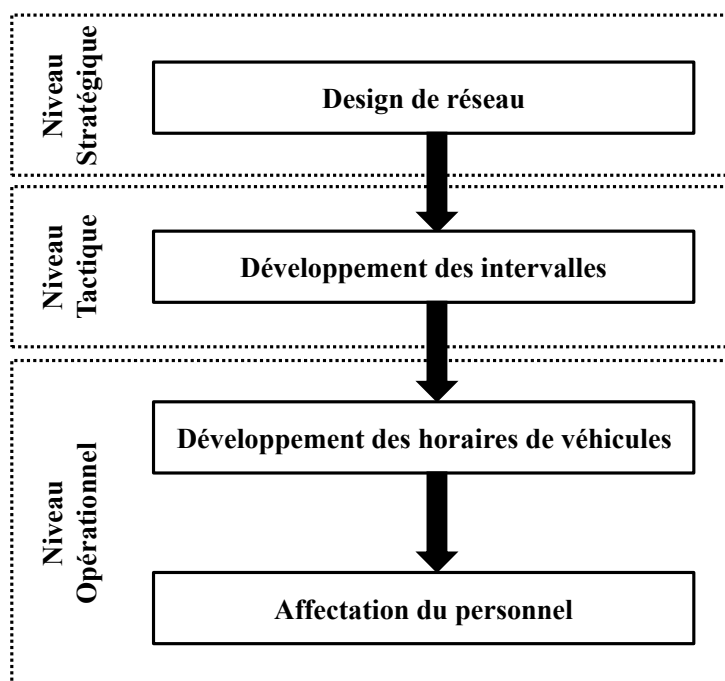


Figure 2.5 Schéma de la POTC en trois niveaux d'étude

(Ceder, 2007; Pelletier et al., 2009)

Dans ce chapitre, de multiples types de données en lien avec les approches de modélisation ont été également identifiés. Le Tableau 2.2 présente une synthèse des avantages ou du potentiel d'utilisation, et des inconvénients ou des limites de chaque type de données. Vu les limites relatives à chaque type de données, certains chercheurs s'intéressent à l'intégration et à la fusion des données (Chapleau & Allard, 2010; Faouzi, Leung, & Kurian, 2011; McCord et al., 2010; Trépanier, Morency, Chapleau, & Blanchette, 2008). La disponibilité des grands ensembles de données et de « *big data* » ouvre de nouvelles voies quant à l'analyse descriptive et de manière plus avancée à la modélisation en TC. Les méthodologies de « *data mining* » peuvent être également bénéfiques en raison de quantités de données importantes dans certains cas. Le « *data mining* » permet de développer un nouveau regard sur l'analyse de données sans avoir fixé un modèle a priori (Hato & Kitamura, 2008).

La modélisation de TC provoque plusieurs types d'erreur dont celles relatives à la construction théorique, à la production des données requises, à la formalisation mathématique et à la calibration des modèles (Bonnel, 2004).

La synthèse des points forts et des points faibles des approches de modélisation présentées au Tableau 2.1 amène certaines réflexions et permet d'identifier des pistes d'améliorations potentielles à la modélisation des déplacements en TC ; c'est sur cette base que s'articulent les objectifs de la recherche. La revue permet en effet de constater plusieurs limites des modèles actuels dont le manque d'aspect comportemental, la complexité technique d'application, la nature séquentielle et la difficulté de transfert dans le temps de certains modèles. Cette thèse vise à contribuer à la modélisation du TC en traitant certaines de ces limites.

Le prochain chapitre consiste à expliquer la démarche de l'ensemble du travail autrement dit l'organisation générale de la thèse en lien avec les limites observées afin d'atteindre nos objectifs de recherche.

Tableau 2.1 Résumé des points forts et faibles des approches de modélisation en transport

<i>Approche</i>	<i>Points forts</i>	<i>Points faibles</i>
Agrégée	<ul style="list-style-type: none"> <li>• besoin de données de référence moins détaillées</li> </ul>	<ul style="list-style-type: none"> <li>• analyse et résultats basées sur un système zonal et pas des données individuelles</li> <li>• structure séquentielle dans la plupart des modèles présentés</li> <li>• formalisation mathématique de certains modèles (ex. forme linéaire du modèle de génération des déplacements)</li> <li>• faible base comportementale dans certains modèles comme le modèle gravitaire</li> <li>• enjeux relatifs à l'estimation des coûts généralisés</li> </ul>
Désagrégée	<ul style="list-style-type: none"> <li>• possibilité d'agréger a posteriori des résultats selon l'objet d'étude</li> <li>• meilleures interprétations comportementales</li> </ul>	<ul style="list-style-type: none"> <li>• besoin de données plus détaillées</li> <li>• plus complexe au niveau des calculs</li> </ul>
Descriptive	<ul style="list-style-type: none"> <li>• base comportementale forte</li> <li>• pratique pour réaliser des analyses de performance en TC</li> </ul>	<ul style="list-style-type: none"> <li>• complexité d'étude au niveau technique</li> <li>• besoin de données important</li> <li>• difficulté de transfert dans le temps</li> </ul>
De simulation	<ul style="list-style-type: none"> <li>• pratique pour évaluer la performance des solutions présentées par d'autres approches</li> <li>• possibilité de produire des données de référence</li> <li>• possibilité d'expérimenter des scénarios divers pour les données de référence</li> </ul>	<ul style="list-style-type: none"> <li>• cumul d'erreurs dues à la nature de l'approche</li> <li>• manque de convergence dans certains cas</li> <li>• sensibilité élevée aux données utilisées</li> </ul>
Opérationnelle	<ul style="list-style-type: none"> <li>• possibilité de produire d'informations pour les voyageurs à partir des résultats obtenus</li> </ul>	<ul style="list-style-type: none"> <li>• complexité de calcul dans certains cas</li> </ul>

Tableau 2.2 Résumé des points forts et faibles relatifs aux données en transport

<i>Données</i>	<i>Points forts</i>	<i>Points faibles</i>
Recensement et inventaires d'emplois et de places scolaires	<ul style="list-style-type: none"> <li>• niveau de résolution fin dans certains cas comme le recensement canadien</li> </ul>	<ul style="list-style-type: none"> <li>• non disponibilité des données à un niveau de résolution souhaité dans certains cas</li> </ul>
Enquête OD	<ul style="list-style-type: none"> <li>• source d'informations détaillées sur les ménages, les personnes et les déplacements</li> </ul>	<ul style="list-style-type: none"> <li>• qualité des données fonction de la mémoire des répondants</li> <li>• manque de continuité dans le temps dans la plupart des cas</li> <li>• divers problèmes relatifs à la qualité des données comme le biais du répondant, etc.</li> </ul>
Enquête à bord	<ul style="list-style-type: none"> <li>• utilisable pour déterminer des matrices de référence OD</li> <li>• bénéfique pour évaluer la sensibilité des voyageurs aux changements dans l'offre du TC</li> </ul>	<ul style="list-style-type: none"> <li>• coûteuse et difficile à réaliser</li> <li>• manque de continuité dans le temps et l'espace</li> </ul>
Données de services planifiés (GTFS)	<ul style="list-style-type: none"> <li>• données détaillées et continues dans le temps et l'espace sur les horaires et la géolocalisation des attributs de services en TC</li> <li>• publiquement disponibles et mis à jour en continu dans la plupart des cas</li> <li>• format unique standardisé permettant la réutilisation des modules développés pour des cas différents</li> </ul>	<ul style="list-style-type: none"> <li>• complexité de traitement relative à la quantité de données importante</li> <li>• difficulté d'intégrer dans certains modèles et outils de modélisation actuels</li> </ul>
Localisation des véhicules de TC (AVL, GPS)	<ul style="list-style-type: none"> <li>• utilisable pour évaluer la ponctualité dans un réseau de TC</li> </ul>	<ul style="list-style-type: none"> <li>• mauvaise qualité de données dans certains cas</li> </ul>

Tableau 2.2 Résumé des points forts et faibles relatifs aux données en transport (suite)

<i>Données</i>	<i>Points forts</i>	<i>Points faibles</i>
Compte à bord	<ul style="list-style-type: none"> <li>• utilisable pour déterminer des matrices de référence OD</li> <li>• représentation de profils de charge dans le temps et l'espace dans le cas de collecte de données automatique</li> </ul>	<ul style="list-style-type: none"> <li>• mauvaise qualité de données dans certains cas</li> </ul>
Systèmes de paiement par CAP	<ul style="list-style-type: none"> <li>• collecte de données automatique et précise de façon continue dans le temps et l'espace</li> <li>• utilisable pour déterminer des matrices de référence OD</li> </ul>	<ul style="list-style-type: none"> <li>• difficile à traiter à cause de l'importante quantité de données</li> <li>• manque de certains attributs tels que le motif, le point de débarquement, etc.</li> </ul>

## **CHAPITRE 3 : DÉMARCHE DE L'ENSEMBLE DU TRAVAIL**

L'objectif de cette recherche est d'améliorer le processus d'analyse descriptive et de modélisation en TC. L'existence de multiples types d'erreurs mentionnés à la fin du Chapitre 2, ainsi que les limites relatives aux différentes approches de modélisation du TC présentées au Tableau 2.1 justifient le besoin d'améliorer le processus de la modélisation en TC. Cette thèse s'articule autour de plusieurs pistes d'amélioration présentées ci-dessous.

### **3.1 Considération des caractéristiques sociodémographiques de la population dans la modélisation de leur comportement de choix d'itinéraires en TC**

La modélisation de choix d'itinéraire en TC dans un réseau complexe urbain de TC est d'un grand intérêt pour les planificateurs de transport. Il nous apparaît intuitivement pertinent de considérer les caractéristiques sociodémographiques des personnes dans la procédure de choix d'itinéraire pour réaliser un déplacement OD en TC. Cela est en lien avec les préférences de chaque personne pour choisir un itinéraire composé de plusieurs éléments tels que les temps d'accès, d'attente et de parcours ainsi que le nombre de correspondances. Nous supposons que la perception de chaque élément dans la structure de choix dépend des préférences de chaque personne selon ses caractéristiques sociodémographiques.

Cette problématique est traitée dans notre premier article intitulé « Demographic Analysis of Route Choice for Public Transit », publié dans la revue « Transportation Research Record » en 2011, et présenté dans le Chapitre 4 de cette thèse (Nazem, Trépanier, & Morency, 2011). En résumé, l'analyse des choix d'itinéraires en TC pendant la période de pointe du matin (6h00-9h00) sur l'île de Montréal (Québec) avec prise en compte des caractéristiques sociodémographiques des personnes, a permis de faire une segmentation sociodémographique basée sur les cohortes d'âge et le sexe. Par la suite nous appliquons un modèle Logit conditionnel dans lequel chaque alternative de choix est représentée par une liste de variables indépendantes décrivant un itinéraire en TC. L'ensemble des choix pour chaque paire OD est produit à partir du modèle Z (Le Gallou, 2009). Ensuite les modèles de choix d'itinéraires par segment de population ainsi que pour la population dans son ensemble sont calibrés.

Les résultats confirment que différents segments de population avec des caractéristiques sociodémographiques différentes font des choix d'itinéraires différents, alors que dans les modèles couramment utilisés la population est traitée comme un ensemble monolithique. Ces résultats montrent que la considération des groupes sociodémographiques permet de mieux modéliser le comportement des personnes en lien avec leurs préférences quant au choix d'itinéraire en TC.

La section 7.1 présente des améliorations apportées à ce modèle en supposant les temps d'accès différents pour les segments de population. Ces temps sont calculés avec les vitesses de marche personnalisées selon les caractéristiques sociodémographiques de chaque segment. Cette approche a permis d'améliorer la performance des modèles de choix d'itinéraires et de mieux représenter le comportement des usagers notamment les personnes âgées dont la vitesse de marche est inférieure aux autres segments de la population.

Les résultats obtenus pourront être bénéfiques en lien avec les enjeux futurs relatifs aux changements de structure dans la composition démographique de la population dans les pays développés notamment le phénomène de vieillissement de la population. Cette nouvelle considération permet de développer des modèles de choix d'itinéraires en TC qui représentent mieux la réalité tout en améliorant la construction théorique et la formalisation mathématique du problème.

### **3.2 Développement d'un modèle intégré de génération-distribution des déplacements en TC à l'aide d'un modèle d'opportunités**

Le développement d'un modèle d'opportunités pour la génération-distribution des déplacements en TC est présenté dans le Chapitre 5 sous la forme d'un article intitulé « Integrated Intervening Opportunities Model for Public Transit Trip Generation-Distribution: A Supply-dependent Approach », publié dans la revue « Transportation Research Record » en 2013 (Nazem, Trépanier, & Morency, 2013).

La raison principale pour laquelle nous examinons le modèle d'opportunités pour la distribution des déplacements en TC est qu'il a des bases davantage comportementales que les modèles de type gravitaire (Kermanshah, 2004). De plus le modèle d'opportunités par rapport au modèle gravitaire est moins sensible à la taille du territoire d'étude (Chow, Zhao, Li, & Li, 2005), et il produit de meilleurs résultats dans les cas où les destinations qui satisfont le motif de

déplacement ne sont pas distribuées d'une manière homogène (Veenstra, Thomas, & Tutert, 2010). Le modèle développé tient compte également des caractéristiques sociodémographiques et socioéconomiques ainsi que l'offre de TC.

Nous avons examiné deux formalisations du modèle d'opportunités. L'objectif est de trouver la formalisation la plus performante afin de modéliser la distribution des déplacements tous motifs en TC. Après avoir segmenté tous les déplacements en trois groupes par motif (travail, étude et autre), les modèles sont séparément calibrés. Les sections 3.2.1 et 3.2.2 résument les deux approches fondamentale et hiérarchique de modèle d'opportunités (Nazem et al., 2012). Également nous proposons sommairement une approche intégrée dans la section 3.2.3. La comparaison de performances des trois modèles (fondamental, hiérarchique et intégré) présentée au Chapitre 7 confirme une meilleure représentation de la réalité en formalisant le modèle d'opportunités d'une façon intégrée, d'où l'idée de développement du modèle intégré de génération-distribution présenté au Chapitre 5.

Les trois sous-sections suivantes présentent les différentes approches examinées pour la formalisation du modèle d'opportunités.

### **3.2.1 Modèle fondamental d'opportunités (MFO)**

Le MFO est un modèle de distribution des déplacements tous motifs. Pour le développement de ce modèle, d'abord trois modèles non linéaires représentant les trois motifs sont indépendamment calibrés. Le résultat de cette étape est le nombre de déplacement de chaque motif pour toutes les paires OD. Par la suite le nombre de déplacements tous motif pour chaque paire OD est calculé en faisant la somme linéaire de tous les déplacements en tout motif pour cette paire OD (Nazem et al., 2012).

### **3.2.2 Modèle hiérarchique d'opportunités (MHO)**

Le MHO est un modèle de distribution des déplacements tous motifs. Ce modèle est calibré à deux niveaux de manière hiérarchique. Le premier niveau vise à calibrer trois modèles non linéaires pour les trois motifs d'une manière indépendante, similaire à ce qu'on a fait pour le développement du MFO dans la première étape. Après avoir calibré ces modèles non linéaires, au deuxième niveau on définit une équation linéaire pondérée qui présente le nombre total de déplacements pour chaque paire OD à partir d'une somme pondérée de tous les déplacements en



tout motif pour cette paire OD. La calibration du deuxième niveau est une calibration linéaire qui donne le poids de chaque motif (Nazem et al., 2012).

### 3.2.3 Modèle intégré d'opportunités (MIO)

Le MIO est un modèle de distribution des déplacements tous motifs. Pour le développement de ce modèle, le nombre total de déplacements pour chaque paire OD est présenté avec une équation non linéaire qui intègre trois équations non linéaires chacune pour chaque motif, et aussi le poids de chaque motif de déplacements. La calibration de ce modèle est une calibration non linéaire qui produit les paramètres ainsi que le poids relatifs à chaque motif (voir la section 0 pour les détails).

Le Chapitre 5 consiste à développer un modèle intégré de génération-distribution, en considérant l'offre de TC par le biais des variables de niveau de service de TC ainsi que certaines variables sociodémographiques (population par trois cohortes d'âge) et socioéconomiques (revenu moyen des personnes). Ce modèle intégré pourra être bénéfique en tant qu'outil d'étude des comportements de mobilité en TC en produisant une matrice OD en fonction des changements apportés au niveau de service du TC et à la structure sociodémographique et socioéconomique de la population. Les variables indépendantes étant prévisibles ou supposables dans le temps, ce modèle peut être utilisé pour prévoir les matrices OD pour le futur. Dans la pratique cela pourra être bénéfique pour certaines études de mobilité à long terme telle que l'ensemencement pour estimer la demande générée lors de nouveaux projets de TC.

Dans ce modèle, le niveau de service de TC dans chaque zone d'étude est représenté par deux variables calculées à partir des données GTFS : le nombre total de passage-arrêts des véhicules de TC durant la journée et la densité spatiale des arrêts dans la zone. Ces deux variables permettent de considérer les changements de natures différentes apportées à l'offre de TC.

Le caractère intégré de ce modèle est avantageux car il s'agit d'une approche non-séquentielle pour la génération et la distribution des déplacements en TC. Par ailleurs les résultats présentés dans le Chapitre 5 démontrent une bonne reproduction du nombre de déplacements produits et attirés.

Un autre caractère du modèle d'opportunités est que le temps de parcours en TC y est utilisé en tant qu'élément implicite pour faire l'ordonnancement des zones de destination à partir de chaque zone d'origine, ce qui fait en sorte que ce modèle est moins sensible à la précision du calcul de

temps de parcours en TC. Autrement dit, tant que l'ordre (selon le temps de déplacement) des zones de destination à partir de chaque zone d'origine ne change pas, il n'y a pas d'effet sur les résultats obtenus du modèle. Par contre ce même caractère pourra rendre le modèle très sensible dans les cas où un changement mineur de temps de parcours modifie l'ordre dans l'ordonnancement des zones. La procédure de choix de destination améliorée présentée dans le Chapitre 6 pourrait être avantageuse face à cette sensibilité.

Les résultats présentés au Chapitre 5 s'inscrivent dans l'optique de l'amélioration des modèles de génération-distribution des déplacements par des bonifications dans leur construction théorique et leur formalisation mathématique. Également, les variables indépendantes utilisées dans ce modèle telles que les caractéristiques sociodémographiques et socioéconomiques de la population ainsi que l'offre de TC le dévoile comme étant un outil pertinent de prévision de la demande en TC pour le futur.

### **3.3 Développement d'une procédure de choix de destination de nature choix discret pour l'utilisation dans un modèle d'opportunités**

Le Chapitre 6 vise le développement d'une procédure améliorée d'ordonnancement des zones de destinations afin de l'insérer dans un modèle d'opportunités pour la génération-distribution des déplacements en TC. Les résultats sont présentés dans l'article intitulé « Revisiting the Destination Ranking Procedure in Development of an Intervening Opportunities Model for Public Transit Trip Distribution », soumis pour publication dans le « Journal of Geographical Systems » en 2013, révisé en 2014 et resoumis. Cela présente en quelque sorte la suite du modèle intégré de génération-distribution développé dans le Chapitre 5. Plusieurs limites du modèle développé au Chapitre 5 sont ciblées et améliorées par le biais de la nouvelle procédure d'ordonnancement des destinations présentée à la section 6.4.3.1 :

- La procédure d'ordonnancement des destinations utilisée dans le Chapitre 5 pourrait devenir très sensible aux changements mineurs dans le temps de parcours en TC. Cela représente les cas où un changement mineur de temps de parcours en TC change l'ordonnancement des zones de destination. Par contre la nouvelle procédure présentée dans le Chapitre 6 n'est pas aussi sensible aux changements mineurs dans le temps de

parcours en TC puisqu'elle utilise plusieurs variables et un modèle de choix discret pour des fins d'ordonnement.

- La procédure d'ordonnement des destinations présentée dans le Chapitre 5 n'est pas sensible au motif et elle donne le même ordre des zones de destination à partir de chaque zone d'origine, peu importe le motif de déplacement. Pour améliorer ce problème dans la construction théorique du modèle, la procédure améliorée d'ordonnement des destinations présentée au Chapitre 6 comprend un ensemble de variables indépendantes qui la rend sensible au motif de déplacement.

Dans l'article présenté au Chapitre 6, nous considérons seulement les déplacements motif travail pour la complexité de préparation des données et de calcul ; rappelons que la méthodologie est généralisable sur d'autres motifs de déplacement. Cet article présente plusieurs améliorations au niveau de la construction théorique et de la formalisation mathématique des modèles de choix de destination et plus globalement du modèle d'opportunités pour la génération-distribution des déplacements en TC.

Les trois Chapitres suivants présentent les trois articles publiés ou soumis pour publication dans les revues scientifiques. Sachant que chaque article est présenté dans son ensemble tel qu'il apparaît dans la revue scientifique, les Chapitres 4, 5, et 6 ont chacun une structure indépendante contenant généralement un résumé, une introduction, une revue de littérature, les données requises, le développement des modèles ainsi qu'une conclusion.

La thèse se termine par une discussion générale et la conclusion.

## CHAPITRE 4 : ARTICLE 1 : « DEMOGRAPHIC ANALYSIS OF ROUTE CHOICE FOR PUBLIC TRANSIT »

rédigé par :

Mohsen Nazem<sup>1</sup>, Martin Trépanier<sup>2</sup>, et Catherine Morency<sup>3</sup>

publié dans :

Transportation Research Record: Journal of the Transportation Research Board, No. 2217(2)  
Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 71-78

DOI: 10.3141/2217-09

The Public Transportation Planning and Development Committee peer-reviewed this paper.

### Abstract

The objective of this study was to understand the behavioral differences between various demographic categories in their choice of public transit route. Until now, almost all models have been done for the population as a whole, but this study attempted categorical choice models. The origin–destination survey in the greater Montreal, Quebec, Canada area was used. For each trip declared, a route choice set was made and then a descriptive analysis was performed. A discrete choice model was applied for six demographic profiles, made up of two genders and three age cohorts. The results showed different coefficients for various profiles. The categorical models could better predict the behavior of individuals compared with the complete model that treated the whole population. Given the issues of aging populations in developed countries, the findings provide a new and vast insight into the future of modeling route choice for public transit.

---

1 M. Nazem, Ph.D. Candidate. Department of Mathematical and Industrial Engineering, École Polytechnique de Montréal. C.P. 6079, Station Centre-Ville, Montreal, Quebec H3C 3A7, Canada. (mohsen.nazem@polymtl.ca)

2 M. Trépanier (Corresponding Author), Professor. Department of Mathematical and Industrial Engineering, École Polytechnique de Montréal. C.P. 6079, Station Centre-Ville, Montreal, Quebec H3C 3A7, Canada. (mtrépanier@polymtl.ca)

3 C. Morency, Associate Professor. Department of Civil, Geological, and Mining Engineering, École Polytechnique de Montréal. C.P. 6079, Station Centre-Ville, Montreal, Quebec H3C 3A7, Canada. (cmorency@polymtl.ca)

## 4.1 Introduction

In modeling, travel or a trip is presented as a movement between two points, origin and destination, with a purpose, by one or several transportation modes and via certain paths called a route. On the basis of the scope of the study, the origin and destination can be presented by points or by zones of different sizes. This study focuses on urban public transit trips, and it aims to better understand the travel pattern of different demographic classes. In other words, in almost all public transit models, behaviors of people belonging to different demographic segments are modeled with a similar hypothesis. This assumption leads the models to treat the population as a whole. In this research, that assumption is put to the test, its impact on predicted behaviors is assessed.

The concept of impedance to travel between two points can be interpreted as a kind of generalized cost or time. It is logically related to the travel time and the transit fare paid by the passenger. In an urban context, most of the time, owing to policies in integrating public transport, travel cost between two points does not depend on the sequence of modes. In such situations, route choice is based just on the utility of paths. A passenger compares different routes in regard to travel time and comfort and chooses the best alternative, on the basis of his or her preferences. Concepts such as travel time and comfort do not have the same meaning for all individuals. One can suppose that in an urban context, for elderly people, the comfort aspects are more important than duration of the travel. In contrast, younger people who are students or workers are less relaxed about duration of travel. These hypotheses led to deepening the focus on public transit route choice preferences of different demographical categories. Owing to demographic structure changes in developed countries, being acquainted with these behavioral differences can help the procedure of modeling at strategic, tactical, and operational levels. This paper will try to demonstrate these differences with a study based on 8 069 observations of the Montreal, Quebec, Canada, 2003 household survey.

The structure of this paper is as follows. First, the methodology is described and then the network and data set used in the study are presented. In the next section, following a descriptive analysis of the variables in the data set, the development of public transit route choice models is done in two parts: the so-called complete model for all passengers as a whole and the categorical models for different population groups. Afterward, results of these models are discussed. Also, for complete appreciation and broader understanding of the issue, some concrete examples of route

choice dissimilarities for users with different demographic characteristics are shown. Then to conclude, some limitations of the present study and opportunities for future work are presented.

## 4.2 Background

The question of route choice in its general form is well studied in the literature (Bonnell, 2004; Hensher & Button, 2008; Ortuzar & Willumsen, 1994), but the problem of route choice in public transit is found less widely in the literature. However, some studies have tackled this question in different ways.

In early research, Hunt presented a Logit model for estimation of public transit route choice (Hunt, 1990). He used some form of generalized cost expression that represents the influence of various factors on the attractiveness of alternatives in a public transport network. He estimated the model using the maximum likelihood procedure described by McFadden (McFadden, 1974). In his study, focus is on work trips at locations in the central business district (CBD) of Edmonton, Alberta, Canada. Hunt mentioned that his approach can be generalized for other trip purposes and locations (Hunt, 1990).

Kato et al. focused on route choice analysis of bus-use commuters in Yangon, Myanmar (Kato et al., 2010). They did their study through a combined revealed preferences–stated preferences approach, by developing a binary-Logit model for the bus route choice. Their study showed the vehicle type preferences of bus-user commuters.

Iseki et al. studied the effects of out-of-vehicle time on travel behavior, especially for transit transfers (Iseki et al., 2006). They developed a transfer penalties framework to evaluate the manner in which different attributes of travel time influence passengers' behavior.

Bovy and Hoogendoorn-Lanser modeled the route choice behavior in a multimodal transportation network using a hierarchical transportation network (Bovy & Hoogendoorn-Lanser, 2005). In their study, “the focus is on preferences for different feeder modes, railway station types and train service types as well as on the relative influence of time elements and transfer penalties.”

Some studies have been done to understand general behavioral differences of various population categories. Newbold et al. studied the travel behaviors within Canada's older population (Newbold, Scott, Spinney, & Kanaroglou, 2005). Their study focused on the number of trips by different modes and its changes within various age cohorts. Although that study is not

exclusively on public transport mode, it showed the effects of demographics on individuals' travel behaviors.

Golob and Hensher tackled the subject in a different way (Golob & Hensher, 2007). They studied trip chaining activities of Sydney, Australia residents by age groups, with a focus on seniors. They also considered all modes in their study. Mercado and Paez focused on distance traveled by elderly people by various modes (Mercado & Paez, 2009). They used Hagerstrand's notions of time geography to understand the ability of individuals to make trips. Krygsman et al. discussed the travel time elements in a multimodal public transport context. They concluded that "access and egress times are a function of access and egress modes and trip orientation rather than of socio-demographic characteristics" (Krygsman et al., 2004).

The review for this paper showed that only a few studies tackled the public transit route choice problem directly for different categories of the population. In other studies, the route choice problem was seen for all the passengers as a whole. Therefore, it is of great interest to formulate this question for different demographic categories, to compare dissimilarities between choices made by these population groups.

### **4.3 Data set**

Mobility data can be used in certain urban planning procedures and in the modeling and characterization of transportation demand. Many other uses for such a survey are presented in the literature (Morency et al., 2006a, 2006b, 2007; Pelletier et al., 2009).

#### **4.3.1 Travel Survey Data**

For almost 40 years, the greater Montreal area has been conducting telephone origin–destination (OD) travel surveys approximately every 5 years (Chapleau et al., 2001). These data include rich information about all trips made by every individual in a 5 % sample of residing households, including precise spatiotemporal details on purposes and all modes used. In 2003, the sample contained almost 400 000 trips. Information such as home location, household size and car ownership, class of income, age, gender and main occupation are also gathered. For this study, 8 069 declared public transit paths were used. These are paths made within the Montreal Island, in the morning peak period.

### 4.3.2 Creation of Alternative Routes Set

To develop a utility-based public transit route choice model, a set of acceptable routes must be simulated for each (OD) pair. For this, the one having maximum utility was identified. The observed paths, those declared in the household survey, will be retrieved in this set of acceptable routes. This step is essential because an individual may not have chosen the route with maximum utility, depending on parameters of the model.

This acceptable route choice set for each OD pair is created automatically with a simulation procedure. First, shortest path trees are computed and stored from every centroid representing an origin in the set of OD pairs. Then, for each OD pair ( $od$ ), the transit network is explored, starting from  $d$ , with a recursive function that builds  $od$  paths. This function assigns a penalty related to a number of statistics about its current  $xd$  path (number of connections, number of arcs that are not part of the shortest paths tree from  $o$ , impedance relative to the shortest path's impedance, etc.) and extends it with an arc entering  $x$  if the resulting path's penalty is smaller than a threshold value. When more than one arc can be used for the extension, they are explored in increasing order of their impedance deficit relative to the shortest path tree. Thus, the arcs in the shortest path tree are always explored first, and the shortest path is the first path generated. In data set here, this process results in several hundred paths for a typical  $od$ . Afterward, a passenger flow repartition model is used (Le Gallou, 2009) to reject most of these paths and keep only the few best ones, those that are assigned a positive flow by this model. This would result in about seven paths for each  $od$ . So, in this stage of the study, a finite set of discrete alternatives was made for each OD trip performed by transit. Because the route chosen by each passenger questioned in the survey was also known, integrating these two data sets according to the common OD pairs results in a new data set that contains the following elements for each pair.

One element is a set of alternative routes and, for each route, all the attributes related to the sequence of modes used (number of bus and subway lines); travel time (access and egress times, waiting times for each mode, and transfer walking times); and transfer type (bus–bus, bus–subway, subway–bus, and subway–subway). The independence of irrelevant alternatives in the choice set is mathematically satisfied. However, if one considers time as is intuitively perceived by individuals, then some choices would become identical; remembering that is this case, one speaks of the intuitive time experienced by public transit users. For example, one can suppose



that two routes of 45 and 47 min are intuitively identical for people, whereas these are not the same in a mathematical choice model.

Another element is the path choice for each person surveyed and all the sociodemographic information related to this person (gender, age, income, purpose, etc.).

For more than 72 % of OD survey registries, the model as described can find the route choice made by the individuals. For about 28 % of the cases for which the model did not find the real route chosen, that alternative was artificially added to the choice set, to estimate the route choice model for all registries in the survey.

## **4.4 Development of Route Choice Model**

In this section, the modeling procedure is described. First, variables are defined, and then a descriptive analysis is performed on attributes related to alternative routes to show the significance level of variability in the data set. To that end, a *t*-test is computed separately for each variable, to understand whether the difference between each two demographic profiles is statistically significant. Results of these *t*-tests are shown in two matrices (Tables 4.1 and 4.2). Afterward, focus is on estimating the two public transit route choice models: the complete model and the categorical model. The former is calibrated for all passengers as a whole, and the latter is estimated separately for different groups of population, that is, so-called demographic categories. The case studied in this paper is a kind of trip-based planning. In the morning peak period, the headways are quite short; therefore, the schedule does not significantly affect the route choice decision of individuals.

Tableau 4.1 t-Test for All Mode and Transfer-Type Variables and for All Demographic Categories

Demographic Categories		0-19 Years Old		20-64 Years Old		65+ Years Old	
		Men	Women	Men	Women	Men	Women
0-19 Years Old	Men: n=1 100						
	Women: n=1 131	NBusBus NBusSub					
20-64 Years Old	Men: n=2 351						
	Women: n=3 370			NBus NBusBus			
65+ Years Old	Men: n=52	NSubBus	NSubBus	NBus NBusBus NBusSub	NBus NBusBus NBusSub		
	Women: n=65			NSubBus	NSubBus	NBus NSub NBusBus NBusSub	

NOTE: Non-significant differences identified.

Tableau 4.2 t-Test for All In-Vehicle and Walking Times and for All Demographic Categories

Demographic Categories		0-19 Years Old		20-64 Years Old		65+ Years Old	
		Men	Women	Men	Women	Men	Women
0-19 Years Old	Men: n=1 100						
	Women: n=1 131	TSub TConWalk TEgrWalk					
20-64 Years Old	Men: n=2 351						
	Women: n=3 370						
65+ Years Old	Men: n=52			TBus TWaitBus TConWalk	TBus TWaitBus TConWalk TEgrWalk		
	Women: n=65		TBus	TBus TEntWalk	TBus TEntWalk TConWalk	TSub TConWalk	

NOTE: Non-significant differences identified.

#### 4.4.1 Definition of Variables

An example of this integrated data set can be seen in Table 4.3. In this data set, each ID represents a record in the transportation survey: in other words, an OD trip made by an individual via public transit. For example, in Table 4.3, one can distinguish three alternative routes for ID 51 and six alternative routes for ID 351591. The chosen alternative for each ID is presented by 1 in the *Choice\_STM* column. Following this column, a set of 15 variables describes the characteristics of each alternative route. At the right of Table 4.3, passengers' sociodemographic attributes, such as gender and age cohort, are presented for each ID. In the data set there are 8 069 IDs that represent the public transit records for weekday trips and morning peak period. Focus was on just the morning peak period because the public transit network is somehow different outside of this temporal period. Hence, it is believed that the current sample is sufficient to test the assumptions.

Variables used in the model are defined as follows:

- NLines = number of lines used (NBus + NSub)
- NBus = number of bus lines used
- NSub = number of subway lines used
- TBus = sum of in-vehicle travel times in buses (min)
- TSub = sum of in-vehicle travel times in subway (min)
- TWaitBus = sum of waiting times for bus [half of headways (min)]
- TWaitSub = sum of waiting times for subway [(half of headways (min))]
- TEntWalk = walking time between centroid of origin zone and first boarding point (min)
- TConWalk = sum of walking times in transfers (min)
- TEgrWalk = walking time between last alighting point and centroid of destination zone (min)
- TWalk = sum of all walking times TEntWalk + TConWalk + TEgrWalk (min)
- NBusBus, NSubSub, NSubBus and NBusSub = number of each transfer type
- Gender = men (1) or women (2)
- Age cohort = 0 to 19 years old (1), 20 to 64 years old (2), or more than 65 years old (3)

The approach is based on a discrete choice modeling procedure for the data set as described. It was decided to apply a conditional Logit model to these data to study the passengers' behaviors and their preferences in regard to utility in the route choice procedure.

Tableau 4.3 Sample of Integrated Data Set Used for Analysis

			<i>Attributes of the Alternative routes (rounded values)</i>															<i>Demographic Characteristics</i>	
<i>ID</i>	<i>Rank</i> <i>Choice_STM</i>		NLines	NBus	NSub	TBus	TSub	TWaitBus	TWaitSub	TWalk	TEntWalk	TConWalk	TEgrWalk	NBusBus	NSubSub	NSubBus	NBusSub	Gender	Age cohort
51	1	1	2	1	1	13	8	3	3	7	5	1	2	0	0	1	0	1	2
	2	0	3	1	2	4	14	4	6	9	5	1	4	0	1	1	0		
	3	0	2	0	2	0	14	0	6	15	5	0	11	0	1	0	0		
351591	1	1	3	1	2	13	6	3	6	7	2	1	5	0	1	0	1	2	3
	2	0	2	1	1	13	5	3	3	9	2	1	7	0	0	0	1		
	3	0	3	1	2	7	6	6	6	9	5	1	5	0	1	0	1		
	4	0	2	1	1	7	5	6	3	11	5	1	7	0	0	0	1		
	5	0	3	1	2	8	4	3	6	13	8	1	5	0	1	0	1		
	6	0	2	1	1	8	3	3	3	15	8	1	7	0	0	0	1		

#### 4.4.2 Descriptive Analysis of Variables

Conducting a descriptive analysis is a priori necessary to understand the variability and pertinence of the data. Two subsets of the integrated data set were analyzed. One subset was all of the routes chosen and declared by passengers in the transportation survey (routes presented with 1 in the *Choice\_STM* column in the sample shown in Table 4.3). The second subset was all of the routes not chosen by passengers (routes presented with 0 in the *Choice\_STM* column in the sample shown in Table 4.3).

The distribution of attributes shows good variability in the variables chosen for the model. The coefficients of variation (*V*) in Table 4.4 confirm that variables that represent the transfer type

(NBusBus, NSub-Bus, and NBusSub) have high variances, whereas other variables have low variances.

Table 4.1 shows results of  $t$ -tests for the variables that represent the number of lines and number and type of various transfers. This table has the form of a matrix for the six demographic categories used in the study. For simplicity, only the variables that are not significantly different are shown. Therefore, when a cell is empty, it means that all variables are significantly different. For example, the variable NBus is not significantly different between the men and the women of 20 to 64 years old. To show the distribution of the data set, the number of observations for each category is shown. These results are extremely useful in interpretation of the models. For example, one cannot highlight any results by comparing the coefficients of the variable NBus between the men and the women of 20 to 64 years old, because these two data sets are not statistically different. In addition, Table 4.2 shows the results of  $t$ -tests for all in-vehicle and walking times.

In the next sections, the choice models are estimated. First, the complete model is estimated for all passengers as a whole. Then the categorical models for different groups of the population are calibrated according to their demographic characteristics.

#### 4.4.3 Estimation of Complete Model for All Passengers as a Whole

The complete model is a Logit model that has been estimated with all the variables shown in the previous section. Some of the variables, such as access times, follow more simplistic assumptions, but they are used anyway in the model. To calibrate the model, the conditional Logit command in STATA is used, which allows estimation of the Logit model for grouped data. The result of the estimation confirms that all variables are significant (all  $p$ -values less than 0,05). The utility function is defined as follows:

$$\begin{aligned}
 \text{Utility} = & -2,728774 * \text{NBus} - 0,5352165 * \text{NSub} + \\
 & 0,8071887 * \text{NBusBus} + 0,2771047 * \text{NSubBus} + 0,6158244 * \text{NBusSub} + \\
 & 0,2007991 * \text{TBus} + 0,2269397 * \text{TSub} + \\
 & 0,1300669 * \text{TWaitBus} - 0,1922711 * \text{TWaitSub} + \\
 & 0,0897885 * \text{TEntWalk} - 0,1270044 * \text{TConWalk} + 0,121417 * \text{TEgrWalk}
 \end{aligned} \tag{4.1}$$

So the probability of choosing a route is this:

$$P(\text{Choosing route } M \text{ in ID } N) = \frac{e^{\text{Utility of route } M}}{\sum_{m \in \text{All routes in ID } N} e^{\text{Utility of route } M}} \quad (4.2)$$

Where  $M$  is the route the probability of which is being calculated, and  $m$  represents all routes in each ID  $N$ .

The complete model can predict about 54 % of the routes chosen by passengers in the survey.

Tableau 4.4 Descriptive Analysis of Variables for Two Data Sets

	Routes Chosen by Passengers Nb. of Observations = 8 069					Routes Not-Chosen by Passengers Nb. of Observations = 42 756				
Variables	Min	Max	Mean	Stand. Dev.	Coef. of Var.	Min	Max	Mean	Stand. Dev.	Coef. of Var.
NBus	0	4	0,93	0,70	76%	0	3	1,37	0,73	53%
NSub	0	3	0,87	0,79	91%	0	3	0,93	0,77	83%
NBusBus	0	3	0,14	0,38	273%	0	2	0,32	0,58	178%
NSubBus	0	1	0,12	0,33	268%	0	1	0,28	0,45	160%
NBusSub	0	1	0,30	0,46	153%	0	1	0,45	0,50	111%
TBus	0	96	14,29	14,85	104%	0	99	15,76	13,76	87%
TSub	0	53	9,68	10,38	107%	0	44	9,13	8,75	96%
TWaitBus	0	44	6,05	6,24	103%	0	43	7,29	5,44	75%
TWaitSub	0	9	2,48	2,15	87%	0	9	2,67	2,08	78%
TEntWalk	1	33	5,26	3,77	72%	1	33	5,09	3,32	65%
TConWalk	0	7	0,57	0,71	124%	0	6	1,09	0,88	80%
TEgrWalk	1	33	5,78	4,29	74%	1	32	5,22	3,71	71%

#### 4.4.4 Estimation of Categorical Model for Demographic Groups

In this section, an attempt is made to find the demographic effects on the utility of various routes. To do so, the choice model is calibrated for different population groups distinguished by gender, age cohort, and the combination of these characteristics.

The dissimilarity in regard to coefficients of the variables in the utility function is presented, while in the next section, some examples of different route choices based on demographic characteristics are given.

Tableau 4.5 Estimation of Utility Function Coefficients by Gender

Variable	Men	Women	Complete Model
NBus	-2,681219	-2,79883	-2,728774
NSub	-0,6943101	-0,3883289	-0,5352165
NBusBus	0,7139552	0,8986162	0,8071887
NSubBus	0,1244305	0,4134516	0,2771047
NBusSub	0,4002887	0,7977089	0,6158244
TBus	0,1978344	0,2037959	0,2007991
TSub	0,2243045	0,2294853	0,2269397
TWaitBus	0,1260038	0,1329127	0,1300669
TWaitSub	-0,1131529	-0,2677089	-0,1922711
TEntWalk	0,0958507	0,083321	0,0897885
TConWalk	-0,063671	-0,1788384	-0,1270044
TEgrWalk	0,1393283	0,1037081	0,121417

##### 4.4.4.1 Gender

The aim is to find out the dissimilarity between the utility function coefficients by gender. According to Table 4.5, which shows these coefficients for male, female, and the complete model, conclusions are as follows:

- Men are more sensitive to the number of subway lines, whereas women are more sensitive to the number of bus lines.
- Women are more sensitive than are men to transfers of all types, especially the transfers between two different modes.



- Men and women have almost the same sensitivity to in-vehicle and waiting times for bus, but women are twice as sensitive as men to the waiting time for subway.
- Men are more sensitive to access times (access and egress), whereas women are more sensitive to walking time in transfers.

#### **4.4.4.2 Age Cohort**

The model is calibrated for population groups, based on three age cohorts, as follows:

- Young people younger than 19 years old,
- Adult people aged from 20 to 64 years old, and
- Elderly people aged 65 years and older.

According to Table 4.6, these points could be highlighted:

- The route choice of the elderly people compared with that of the adult category is less affected by the number of bus lines and the number of transfers of all types.
- Young and adult-aged people have almost the same sensitivity to the number of subway lines, whereas the elderly are about four times more sensitive to this variable than are young and adult-aged people.
- Elderly people tolerate the in-vehicle and waiting times much better than do those in the other age categories.

#### **4.4.4.3 Gender and Age Cohort**

The model for six demographic categories is estimated as described in Table 4.7, to better understand the behavioral differences of each gender in each age cohort. According to Table 4.7, one can see the following:

- The number of bus lines decreases the utility of a route for young women more than it does for young men.
- The effect of the in-vehicle time is almost the same with just minor differences for both genders in the adult category.
- In the elderly category, men are more sensitive than women are to access and egress times.

From Table 4.7, one can see that, for some variables, there is no justifiable tendency. This may stem from the insufficient number of observations for some categories in the modeling procedure.

Tableau 4.6 Estimation of Utility Function Coefficients by Age Cohort

Variable	0-19 Years Old	20-64 Years Old	65+ Years Old	Complete Model
<b>NBus</b>	-1,402815	-3,23278	-1,915241	-2,728774
<b>NSub</b>	-0,4838613	-0,4523281	-2,019336	-0,5352165
<b>NBusBus</b>	-0,3910019	1,213067	0,1018914	0,8071887
<b>NSubBus</b>	-0,5544896	0,5744718	0,1860423	0,2771047
<b>NBusSub</b>	-0,0634672	0,8986403	0,0471964	0,6158244
<b>TBus</b>	0,1994156	0,2026648	0,2711043	0,2007991
<b>TSub</b>	0,2292637	0,2235935	0,2632418	0,2269397
<b>TWaitBus</b>	0,115021	0,1367546	0,2170187	0,1300669
<b>TWaitSub</b>	-0,062054	-0,258827	0,446564	-0,1922711
<b>TEntWalk</b>	0,1074297	0,0786824	0,1810115	0,0897885
<b>TConWalk</b>	-0,0234028	-0,1524951	-0,1303718	-0,1270044
<b>TEgrWalk</b>	0,1494069	0,1087496	0,1058751	0,121417

#### 4.4.5 Examples of Route Choice for Demographic Categories

Even though the categorical models have shown differences in behaviors, these differences could be insufficient to distinguish the paths between origin and destination pairs. In this section, some examples of public transit route choice in the greater Montreal area are demonstrated to confirm the dissimilarity of choices based on demographic categories. To that end, the OD pair was chosen, the route choice was estimated by means of our categorical models, and the estimation was compared with the real route choice declared in the survey.

Figure 4.1 shows four cases of such comparison for various OD pairs and demographic categories. In this figure, two dotted point lines show the routes chosen by men and women according to the categorical models, and the solid line shows the real choice made by a public transit user. For example, Case II shows the trip made by a woman 53 years old, between Plateau Mont Royal and the CBD of Montreal. In this case, if the men's model is used, the chosen route will be Subway 2, whereas the women's model chooses Bus 47 and Subway 2. The real choice made by the public transport user in this case is also Bus 47 and Subway 2, which shows the right

route choice based on the individual gender. However, Case I can partially predict the route choice. In other words, in this case, the categorical model for men predicts a bus–subway combination route, whereas the categorical model for women predicts a subway route. The real choice made by the 25-year-old man is a bus–subway combination route that is more similar to the prediction of the categorical model for men.

The number of bus and subway lines shown in Figure 4.1 does not match the real bus and subway line numbers in Montreal. That is because a different codification system was used with different line numbers.

Tableau 4.7 Estimation of Utility Function Coefficients by Age Cohort and Gender

	0-19 Years Old		20-64 Years Old		65+ Years Old		
Variable	Men	Women	Men	Women	Men	Women	Complete Model
NBus	-0,9714272	-1,816573	-3,445926	-3,115979	-2,0481	-2,24319	-2,728774
NSub	-0,8268907	-0,1602836	-0,5623111	-0,3728873	-4,203275	0,005573	-0,5352165
NBusBus	-0,6986453	-0,1152737	1,225646	1,207599	0,3072987	0,489874	0,8071887
NSubBus	-0,6787104	-0,4297112	0,4753285	0,6570959	-0,3567391	0,8806782	0,2771047
NBusSub	-0,5482185	0,3812289	0,8453791	0,9529826	-0,5801833	1,055115	0,6158244
TBus	0,1957677	0,2050124	0,2025267	0,2034871	0,2680756	0,2678773	0,2007991
TSub	0,2353965	0,22651	0,2175448	0,227855	0,2362564	0,2783965	0,2269397
TWaitBus	0,1068361	0,1228971	0,1358103	0,1370447	0,2829336	0,1670261	0,1300669
TWaitSub	0,0995224	-0,2217964	-0,2119565	-0,2970768	1,56411	-0,659358	-0,1922711
TEntWalk	0,1068524	0,1088486	0,085682	0,0705687	0,1840914	0,1392619	0,0897885
TConWalk	-0,0465017	-0,0055945	-0,0459716	-0,2347712	-0,6280404	-0,0120924	-0,1270044
TEgrWalk	0,1611432	0,1376676	0,1285768	0,0901122	0,1130472	0,0862877	0,121417

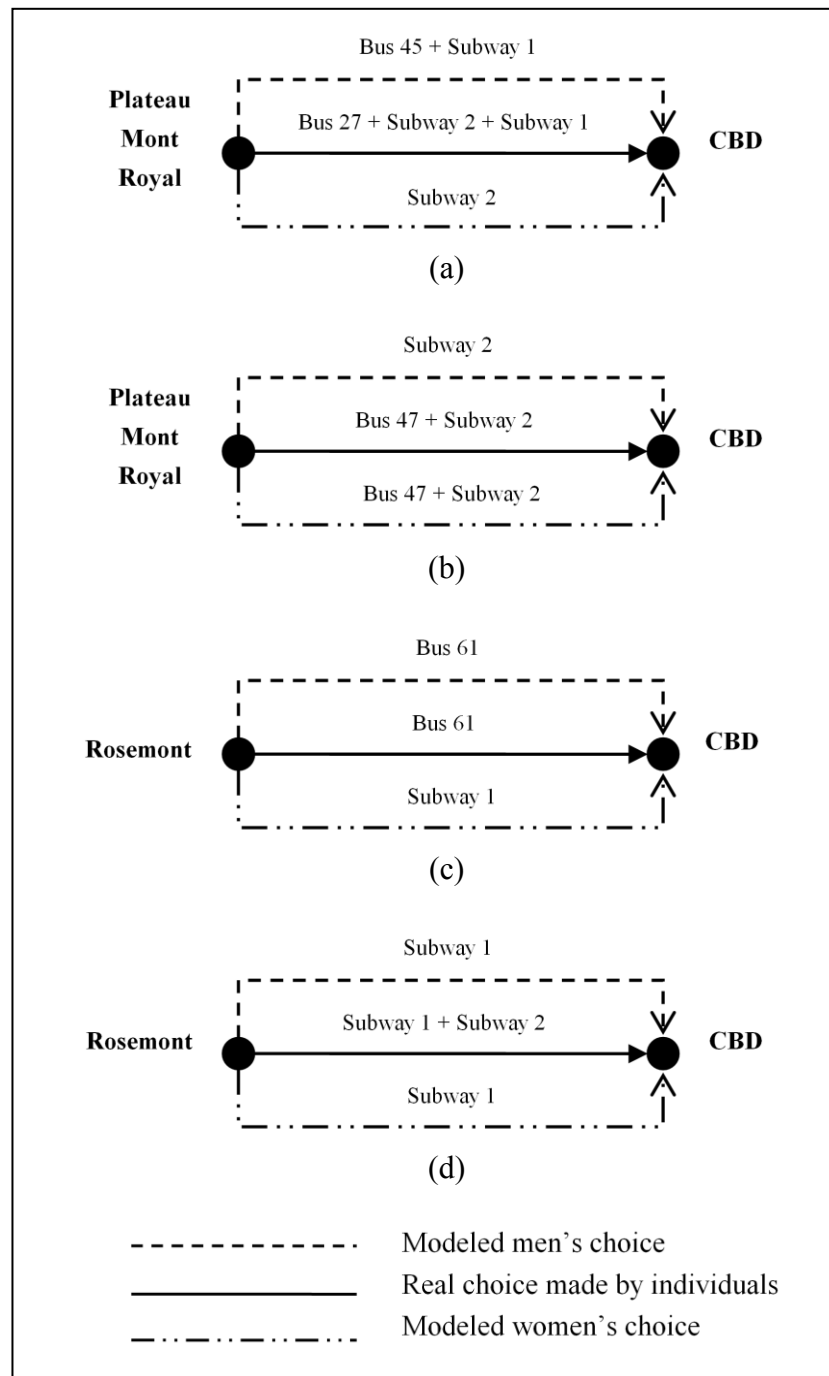


Figure 4.1 Examples of Route Choice in the Greater Montreal Area

(a) Case I, 25-Year-Old Man; (b) Case II, 53-Year-Old Woman;  
(c) Case III, 51-Year-Old Man; and (d) Case IV, 35-Year-Old Woman.

## 4.5 Conclusion and Perspectives

In this study, the problem of route choice in an urban public transit network was addressed in a different manner. In almost all previous studies, the population is treated as a whole, with no consideration about the preferences of various demographic categories. This research showed that the route choice behavior of these population groups can be different. To that end, different kinds of discrete choice models were developed for the population as a whole. Then the choice models for six demographic profiles, based on two genders and three age cohorts, were estimated. Use of a complete form of utility function was attempted, including mode types, transfer types, and components of times from an origin to a destination. The model estimations showed that the utility functions for different profiles have various coefficients for the same attribute. This confirms the hypothesis that the behavior of different demographic categories is not the same. Throughout the last part of this paper, some examples were shown of route choices made by the categorical models developed. These choices were compared with real choices made by public transport users. This comparison showed that the categorical models are more efficient. In fact, for more than 600 of 8 000 cases, the demographic categorical models predict the real choice made by individuals; but in these cases, the complete model is not able to give the right route choice. In about 8 % of the cases, the categorical model, compared with the complete model, gives better results; in other cases, the complete model can predict the real choice made by individuals. In other words, this approach can improve the global accuracy of models of public transit route choice, by improving the accuracy for some IDs.

These findings show a real need for disaggregating the route choice models in public transit, to better estimate the behavior of various demographic groups.

This study made the first step in this domain. For future research, some improvements are proposed in the form of the utility function. To that end, the use of different variables for the walking times related to various types of transfer is suggested. With this improvement, one can better distinguish the unattractiveness related to each type of transfer, in regard to walking time.

In this study, walking distances are calculated in a particular way because of some computational complexity in the procedure of creating the route choice set. It can be an interesting improvement for future research, to calculate the exact access (egress) times, in accordance with the real dwelling (destination) places.

Another proposition for future research pertains to the use of a nested Logit model. Three nests could be used, such as mode type, transfer type, and components of times. However, in that case, attention should be paid to the conditions of use of nested Logit models.

An improvement can be made in this research by substituting walking times with walking distances. If that were done, one could assign different walking speeds for various demographic categories, especially for different age cohorts, to estimate walking time in its most realistic form. The calibration of the model then could be done with these times.

In this study, trips done during the morning peak hours were analyzed. Although this presents a sort of limitation, the methodology remains the same in the case of an entire day. The number of observations for each profile shows that the data set is biased toward individuals from 20 to 64 years old. The number of data points in the third age group (more than 65 years old) is little; this is probably the reason for the statistical similarity between many variables presented in Tables 4.1 and 4.2. This can be an interesting area for future research, to extend the model on an entire day. By calibrating the models for an entire day, more realistic and reliable results for all demographic groups could be obtained, because the number of observations for different groups will be closer to uniform.

As seen, the topic of public transit route choice is vast, and the model estimations can be done for various subgroups of population, to better estimate the behavior of public transport users. With current issues, such as aging of the population in developed countries, these realistic considerations in the planning procedure can help public transit organizations toward more sustainable transportation.

## **Acknowledgments**

The authors acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada. They give special thanks to Jonathan Dumas for developing the route choice sets. The authors also acknowledge the Agence Métropolitaine de Transport and the survey consortium that provided the origin–destination survey data.

# CHAPITRE 5 : ARTICLE 2 : « INTEGRATED INTERVENING OPPORTUNITIES MODEL FOR PUBLIC TRANSIT TRIP GENERATION-DISTRIBUTION, A SUPPLY-DEPENDENT APPROACH »

rédigé par :

Mohsen Nazem<sup>4</sup>, Martin Trépanier<sup>5</sup>, et Catherine Morency<sup>6</sup>

publié dans :

Transportation Research Record, Journal of the Transportation Research Board, No. 2350 (1)  
Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 47-57  
DOI: 10.3141/2350-06

The Public Transportation Planning and Development Committee peer-reviewed this paper.

## Abstract

An Integrated Intervening Opportunities Model (IIOM) was developed for Public Transit (PT) trips. This model is generation-distribution and supply-dependent, with single constraints only on trip production values for work and study PT trips made during morning peak hours (6:00 AM to 9:00 AM) within the Island of Montreal, Quebec, Canada. Several data sets, including the 2008 origin-destination survey of the Greater Montreal Area, 2006 census of Canada, General Transit Feed Specification network data, and school enrollment data, along with the geographical data of the Greater Montreal Area, were used. The IIOM is a nonlinear model with sociodemographic, socioeconomic, and PT supply characteristics, as well as work and study spatial location attributes. Analysis of the modeling performance by means of several Goodness-of-Fit measures showed that the IIOM was well behaved (i.e., globally it had good prediction capabilities) and

---

4 M. Nazem (Corresponding Author), Ph.D. Candidate. Department of Mathematical and Industrial Engineering, École Polytechnique de Montréal. C.P. 6079, Station Centre-Ville, Montreal, Quebec H3C 3A7, Canada. (mohsen.nazem@polymtl.ca)

5 M. Trépanier, Professor. Department of Mathematical and Industrial Engineering, École Polytechnique de Montréal. C.P. 6079, Station Centre-Ville, Montreal, Quebec H3C 3A7, Canada. (mtrepanier@polymtl.ca)

6 C. Morency, Associate Professor. Department of Civil, Geological, and Mining Engineering, École Polytechnique de Montréal. C.P. 6079, Station Centre-Ville, Montreal, Quebec H3C 3A7, Canada. (cmorency@polymtl.ca)

more accurate than the classical gravity model. On the basis of explanatory variables used in the IIOM, the study presents a new tool for PT analysts, planners, and policy makers for studying potential changes in PT trip patterns, as a result of changes in sociodemographic and socioeconomic characteristics, PT supply, and so on. Also, this study opens new opportunities for development of more accurate PT demand models with new emergent data such as smartcard entries.

## 5.1 Introduction

Trip distribution is the second step in the classical sequential four step models for aggregate transport planning. Several families of trip distribution models, such as the Gravity Model (GM) and Intervening Opportunities Model (IOM) are presented in the literature (Bonnell, 2004; Hensher & Button, 2008; Ortuzar & Willumsen, 1994; Wilson, 1970). In the GM, the number of trips between each Origin-Destination (OD) pair is based directly on the OD distance and trip production and attraction values, while the IOM considers the number of intervening opportunities as the main influencing factor. Because it seems that both of these factors influence the trip distribution, some researchers developed unified hybrid models (Almeida & Goncalves, 2001; Goncalves & Uiysséa Neto, 1993; Wills, 1986). In past studies, especially in the case of hybrid models, trip distribution is calibrated usually just for one purpose: work or study. Thus, the advantages and disadvantages of them are not well known.

In this study, the aim is to calibrate an Integrated Intervening Opportunities Model (IIOM) for work and study Public Transit (PT) trips. This model considers the sociodemographic and socioeconomic attributes as well as the PT supply characteristics of Municipal Sectors (MS) located in the Island of Montreal, Canada. In other words, the transportation analysis zones in this study are the MS presented in Figure 5.1. The advantage of this model is that it allows policy makers to study the effects of supply modifications on the PT trip pattern. Also, because the required entry data for this model could be available for the future, it could be used for forecasting purposes.

On the basis of past studies, each family of distribution models has its own advantages and disadvantages compared with other models. At the same time, the calibration of each model requires different data and calibration procedures. The IOM has been used less than the GM has because of complexities in the calibration procedure (Goncalves & De Cursi, 2001). However,



compared with the GM, the IOM is behavioral based (Kermanshah, 2004) and less sensitive to the size and shape of study area (Chow et al., 2005). The IOM also produces better results in cases where destinations that satisfy the trip purpose are not uniformly distributed, like discrete attraction points, such as for shopping or study purposes (Veenstra et al., 2010).

For the calibration of most distribution models, an OD reference matrix is needed. Several approaches are presented in the literature, such as asking passengers to fill out questionnaires on board of transit vehicles, estimating the number of passengers from counting the boarding and alighting passengers at stops, or from smart card validation data (S. Feng & Li, 2004). This study uses the OD matrix derived from the large 2008 OD survey of the Greater Montreal Area (GMA). In previous studies, the IOM was calibrated with two approaches: basic IOM and Hierarchical IOM (HIOM). Results showed that the IOM compared with the GM has much better performance in trip production and trip attraction satisfaction and in other Goodness-of-Fit measures (Nazem et al., 2012). This model can be used later in middle- and long-term forecasting. The models are developed at the MS level, for PT trips done during normal weekday peak hours (6:00 AM to 9:00 AM), within the Island of Montreal. This area consists of 41 MS, with a population of nearly 1,9 million, distributed in almost 500 km<sup>2</sup> (StatCan, 2011b). This study is limited to the Island of Montreal to deal with less calculation complexities, remembering that the methodology can be generalized on the whole GMA.

In contrast with the previous HIOM, with a hierarchical nature (Nazem et al., 2012), in this study, the IIOM is developed and integrated as one nonlinear equation. The total number of trips for each OD is defined by an integrated equation consisting of two distinct nonlinear formulations for work and study trips, and two linear weighting coefficients (Equations 5.3 through 5.5). Afterward, the IIOM is calibrated to calculate the weighting coefficients and the model parameters. A great advantage of the newly developed IIOM compared with the previous HIOM (Nazem et al., 2012) is that all required entry data for development of the new model come from external and independent sources (Figure 5.2). That presents a great predictive capacity for the new IIOM. Also, the sociodemographic, socioeconomic, and PT supply variables allow study of the effect of changes in these variables on the PT trip pattern.

The structure of the paper is as follows. First is a literature review to learn more about advances in trip distribution models; then the data sets and their preparation for modeling development are described. In the next section, following a descriptive analysis of the data, the development of the

IIOM is presented. Afterward, performance of the IIOM is studied and compared with the classical GM by means of several Goodness-of-Fit measures. The next section presents spatial limitations of the IIOM by means of a spatial residual errors analysis. In the conclusion is a discussion of the potential of the IIOM for analysts and policy makers; there is also a presentation of some ongoing and future research topics.

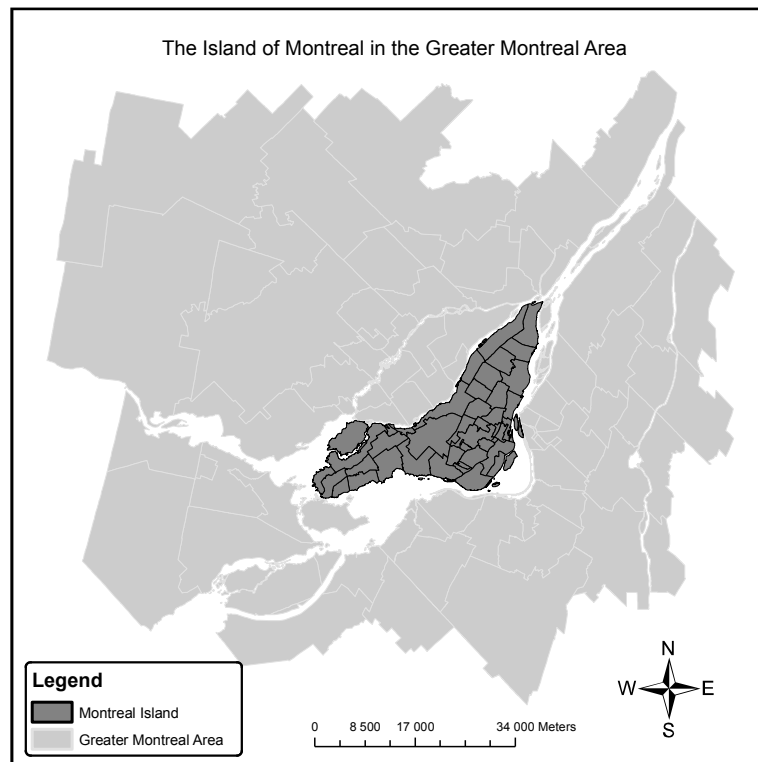


Figure 5.1 Spatial Distribution of the MS on Island of Montreal

## 5.2 Literature Review

Several families of models are used for estimating trip distribution in its general aggregate form. Among them, the GM and the IOM are the most common (Bonnell, 2004; Ortuzar & Willumsen, 1994).

The GM for trip distribution, inspired by Newton's law, uses an impedance function, generally represented by a generalized cost. This model and its applications are well presented in the

literature (Bonnell, 2004; De Grange et al., 2009; Hensher & Button, 2008; Ortuzar & Willumsen, 1994; Rajesson, 2009; Thamizh Arasan et al., 1996; Wilson, 1970).

The main idea of the opportunity model came from some theoretical concepts that relate mobility, migration distances, and spatial locations of services; the theory of this model in its present form was developed later (Schneider, 1959; Stouffer, 1940, 1960). The fundamental idea of this model is that generalized cost is not the only factor that affects destination choices. In contrast, this model considers that the relative accessibility of opportunities that can satisfy trip purpose is the main influencing factor. This model assumes that an individual chooses the closest destination location that gives him/her the opportunity to meet his or her needs. Distance, or more widely generalized cost, is not a continuous variable anymore as it was used explicitly in the GM, and it serves rather to find the ranking of destinations from a given origin point (Bonnell, 2004; Hensher, 1977; Hensher & Button, 2008; Ortuzar & Willumsen, 1994). The use of this model in transportation planning is briefly presented in the literature. In the 1980s, the IOM was used for modeling during the Chicago, Illinois, area transportation studies (Eash, 1983, 1984). More recently, the IOM was used for simulating student flows, and results confirmed that such a model has better performance than the GM (Almeida & Goncalves, 2001). To consider both distance and intervening opportunities in a single trip distribution model, an integrated gravity-intervening opportunities model is also presented and tested in the literature (Goncalves & Ulysséa Neto, 1993; Wills, 1986). In the current study, the aim is to develop and calibrate an IIOM for work and study trips and compare its performance with a classical GM.

For calibrating trip distribution models, reference OD matrices are generally required. The matrices could be estimated by several ways: direct observation, synthesis, and so on (Yaldi, Taylor, & Yue, 2011). In this study, the OD matrices are obtained by processing the data collected during a large-scale OD survey held in 2008 in the GMA (AMT, 2010).

After calibration of trip distribution models, some Goodness-of-Fit measures are needed. Several measures are presented in the literature (Akwawua & Pooler, 2001; Black, 1991; Chow et al., 2005; Evans & Pooler, 1987; Hu & Pooler, 2002; Smith & Hutchinson, 1981; Yaldi et al., 2011). Formulations of required Goodness-of-Fit measures and other related works are presented in the following sections.

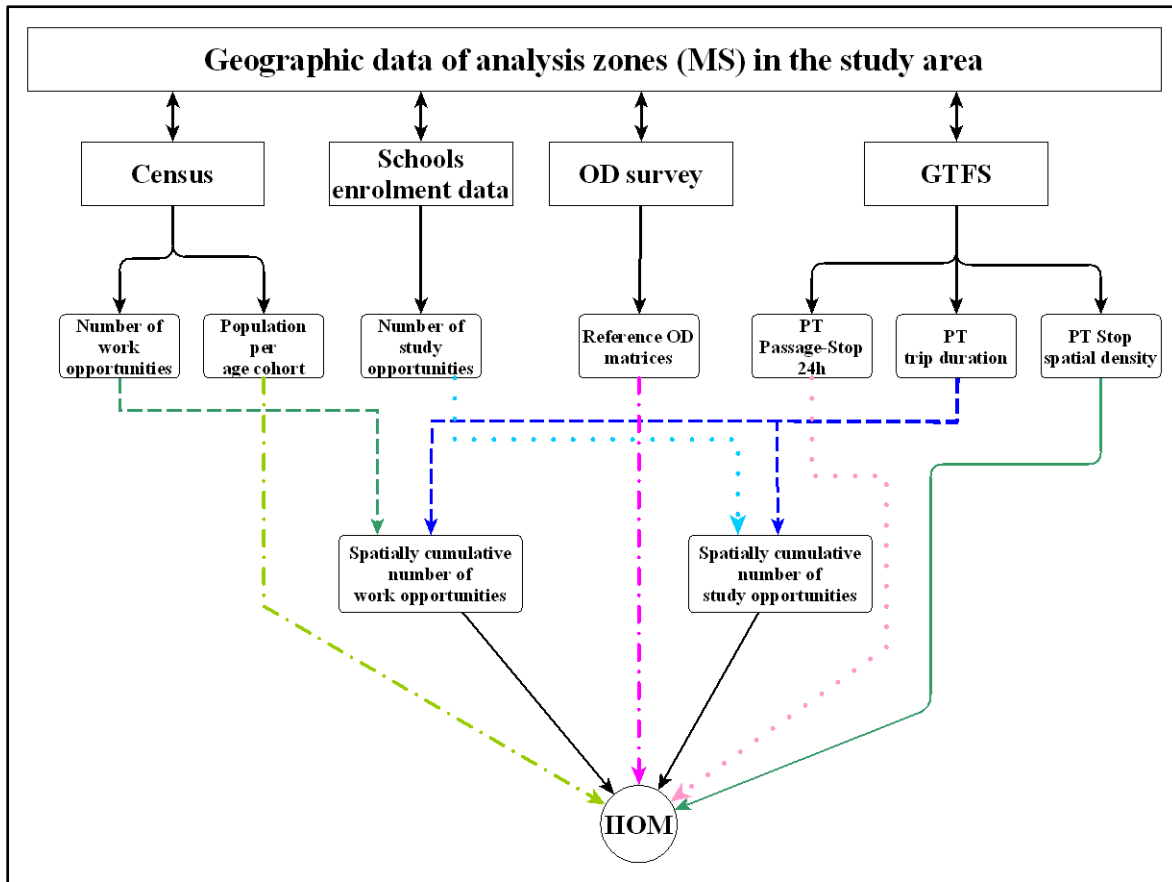


Figure 5.2 Data Flow Diagram for Development of IIOM

### 5.3 Data for Analysis

This section introduces the data sets used in the study. First the data sources are presented; afterward, their preparation for modeling is shown.

#### 5.3.1 Greater Montreal Area OD Survey

For almost 40 years, GMA authorities have been conducting telephone OD travel surveys approximately every 5 years. These data include rich information on all trips made by every person in a 5 % sample of residing households. This makes the OD survey the primary source of information on people's movement habits, according to a report (AMT, 2010). Precise spatiotemporal details are collected on all-purpose and all-mode trips. This study uses data from the most recent OD survey, conducted in 2008. In 2008, the sample contains almost 319 900

trips. Demographic information, such as dwelling location; household size; car ownership; and class of income and age, gender, and main occupation is also gathered. Each record provides an expansion factor that will be used to expand the data set based on the collected 5 % sample (AMT, 2010).

In this study, the OD survey was used for deriving the reference OD matrices for each trip purpose; the process is explained in the section on data preparation.

### **5.3.2 Census of Canada**

Census of Canada is a “unique undertaking on a vast scale” conducted every 5 years by Statistics Canada (StatCan, 2011b). It consists of collecting data from 31,6 million people and more than 13,5 million dwellings. This study used data from one of the most recent censuses, conducted in 2006 (StatCan, 2011a), and it derives the population per age group, number of opportunities for work trips, and individual average income in each MS.

### **5.3.3 School Enrolment Data Set**

The number of opportunities for study trips is derived from the school enrollment data set. This data set lists all students and their related school; it allows estimation of the capacity of educational institutions.

### **5.3.4 General Transit Feed Specifications**

According to the definition given by Google, “the GTFS defines a common format for PT schedules and associated geographic information. GTFS feeds allow PT agencies to publish their transit data and developers to write applications that consume that data in an interoperable way” (Google, 2012). This study used Island of Montreal General Transit Feed Specification (GTFS) obtained from the Société de Transport de Montréal, the PT agency in the Island of Montreal.

The GTFS is used for several purposes, such as characterizing the PT Level-Of-Serve (PTLOS) in each MS, and for calculating the PT trip duration for each OD pair. The PTLOS is represented with two variables that are sensitive to changes in the PT service characteristics, such as headway, and so on. These variables are the total number of Passage-Stops (transit vehicle passing a stop) per 24 h in each MS and the spatial density of stops in each MS.

For calculating PT trip durations for each OD pair, because this study is aggregate at the MS level, the geographic centroid of each MS is considered as its spatial delegate. Morning peak hour weekly schedules from the Montreal GTFS data were used to get the shortest routes and related travel times and to finally obtain a PT travel time matrix among all 41 MS centroids in the Island of Montreal

In an IOM, trip duration is used to calculate the spatially cumulative opportunities for each OD pair. The methodology is described later.

### 5.3.5 Data Preparation

This section presents data preparation for calibrating the IIOM, for the required data sets in modeling. On the basis of the presented data sources, following is a description of creation of reference OD matrices, number of opportunities in each MS, number of spatially cumulative opportunities, and total number of Passage-Stops per 24 h in each MS.

#### 5.3.5.1 Reference OD Matrices

Two OD matrices, one for work trips and the other for study trips are created from the 2008 OD survey data. Each record in the OD survey presents the complete characteristics of an individual movement. First, all return-home and off-peak trips are excluded, to obtain the required data. Afterward, the  $T_{ij}$  values are calculated for each purpose by summing up the expansion factor values for all concerned OD pairs. This results in two distinct data sets: for work trips and for study trips. Since the OD survey is limited to a 5 % household sample, some OD pairs will have zero trips for some trip purposes.

This kind of OD reference data set is usually presented in a matrix form, but in this study, columns were turned into rows to obtain a table form. After calculating the number of trips for each MS, one can obtain trip production and attraction values for each MS and trip purpose. Figure 5.3 presents a sample of deriving trip number ( $T_{ij}$ ), production ( $E_i$ ), and attraction ( $A_j$ ) values for work trips.

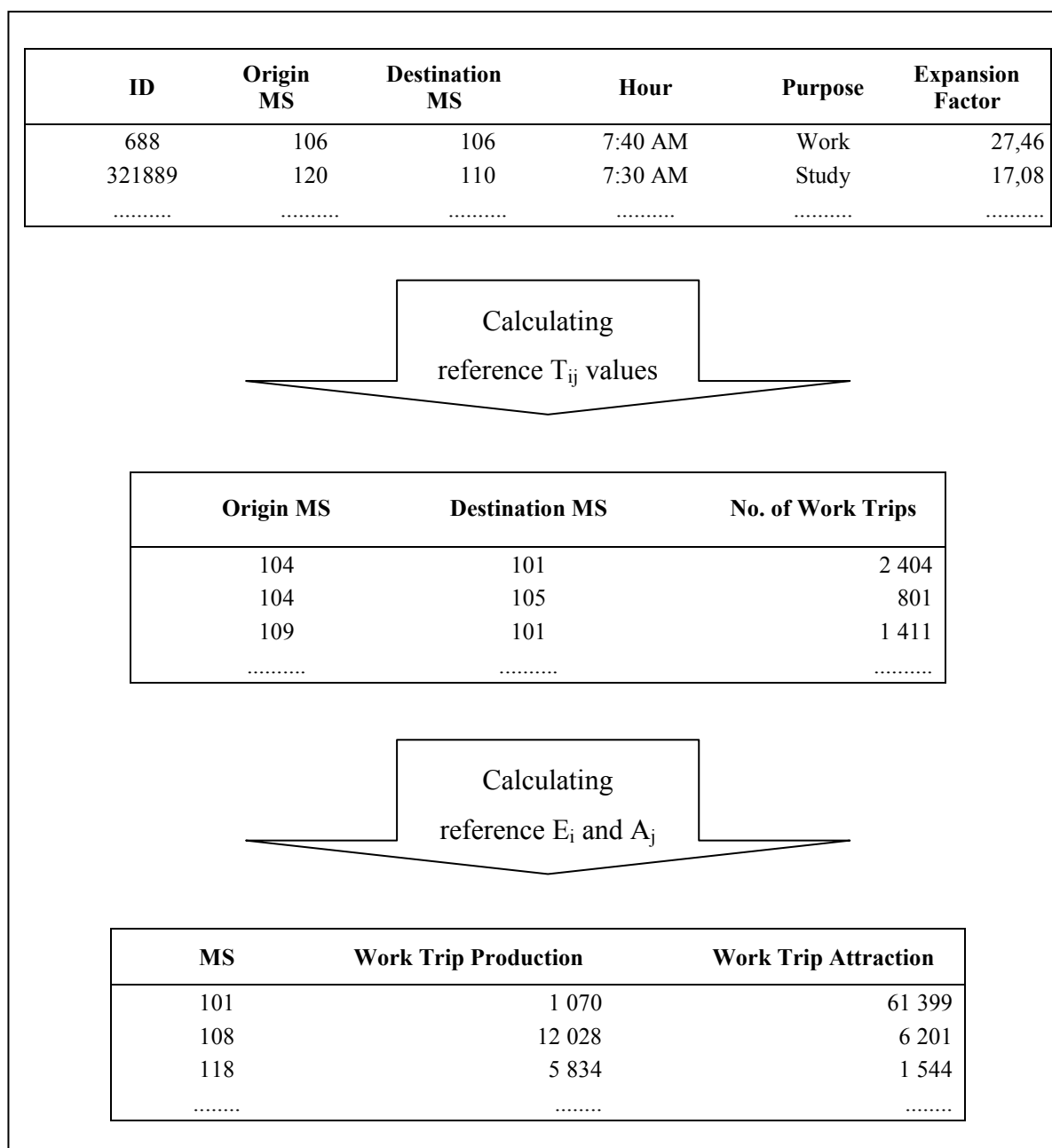


Figure 5.3 Deriving the Reference OD Matrices and Trip Production and Attraction Values from the OD Survey

(Ellipses=Additional Rows; ID=Identifier; No.=Number)

### 5.3.5.2 Work and Study Opportunities in Each Municipal Sector

The development of trip distribution models needs the number of opportunities in each MS. Since the number of opportunities that can satisfy each trip purpose depends on activity locations that can fulfill that activity type, distinct data sets are needed for different trip purposes.

The number of work opportunities in each MS is derived from the 2006 census of Canada data, and the number of study opportunities, except universities, is derived by summing up the estimated capacities of educational institutions in each MS. The spatial distribution of work and study opportunities confirms that the distributions of opportunities for different purposes are not the same, as well as the necessity to consider two distinct purposes for application of the IOM.

### 5.3.5.3 Spatially Cumulative Opportunities for Work and Study Trips

For development of an IOM, one needs the number of intervening opportunities for each OD pair. Because in the study area the PT fare remains constant, only PT trip durations calculated via GTFS are considered to calculate the generalized cost for each OD pair. On the basis of this hypothesis, for each purpose, the number of intervening opportunities between zones  $i$  and  $j$  is the sum of all opportunities located between  $i$  and  $j$ . This number is called “spatially cumulative opportunities,” and it will be used in the calibration of the IOM. For each purpose, this number can be calculated by executing the following steps:

- Step 1: From each origin MS, enumerate all destination MS on the basis of the increasing PT trip times.
- Step 2: Sum up all opportunities in each destination MS to find the total number of opportunities in each MS.
- Step 3: Calculate the number of spatially cumulative opportunities including each destination MS, on the basis of the ranked MS.
- Step 4: Calculate the number of spatially cumulative opportunities excluding each destination MS, on the basis of the ranked MS.

Table 5.1 shows the form of the derived data set, for two origins (MS 101 and 141) and for work trips.

For an idea about the data flow in this study, Figure 5.2 shows main data sources and two levels of data processing for developing the IOM.



## 5.4 Development of Integrated Generation-Distribution Model

In this section, the IIOM is developed: an integrated model of generation–distribution that is supply-dependent. In other words, using variables describing the PT supply in each MS, the model calculates the OD matrices as well as trip production and attraction at the MS level.

Tableau 5.1 Sample of Number of Spatially Cumulative Opportunities for Work Trips

<b>i</b>	<b>j</b>	<b>ij</b>	<b>PT Trip Time in Ascending Order</b>	<b>No. of Work Opportunities in MS j</b>	<b>No. of Spatially Cumulative Work Opportunities between i and j, Including j</b>	<b>No. of Spatially Cumulative Work Opportunities between i and j, Excluding j</b>
101	101	101101	0	182 215	182 215	0
	102	101102	13	106 045	288 260	182 215
	106	101106	15	43 994	332 254	288 260
	107	101107	19	48 878	381 132	332 254
	105	101105	26	65 779	446 911	381 132
	.....					
	138	101138	74	3 165	1 095 350	1 092 185
	137	101137	75	9 780	1 105 130	1 095 350
	140	101140	85	4 485	1 109 615	1 105 130
	134	101134	96	1 835	1 111 450	1 109 615
	141	101141	102	1 165	1 112 615	1 111 450
.....						
141	141	141141	0	1 165	1 165	0
	140	141140	30	4 485	5 650	1 165
	139	141139	49	4 360	10 010	5 650
	137	141137	53	9 780	19 790	10 010
	131	141131	55	33 140	52 930	19 790
	.....					
	117	141117	141	27 245	1 049 580	1 022 335
	116	141116	151	28 665	1 078 245	1 049 580
	113	141113	153	12 135	1 090 380	1 078 245
	114	141114	158	15 275	1 105 655	1 090 380
	115	141115	160	6 960	1 112 615	1 105 655

### 5.4.1 Descriptive Analysis of Data

This study used the 2008 Montreal OD survey data set for all work and study trips done by PT during weekday morning peak hours (6:00 AM to 9:00 AM) within the Island of Montreal. After applying the expansion factors, 278 005 PT trips were obtained for calibrating the model. These a.m. peak hour PT trips within the Island of Montreal represent 67,5 % of all PT trips in the GMA, and more globally, 37,6 % of all 24 h PT trips in the GMA. The distribution of all PT AM peak trips based on purpose is 57,1 % for work, 36,7 % for study, and 6,2 % for other purposes. In this study, the aim is to model work and study trips that gather almost 94 % of all PT AM peak trips.

### 5.4.2 Development of IOM

The IOM in its general form is presented as the following:

$$T_{ij} = \alpha \cdot E_i \cdot (e^{-P \cdot O_{j-1}} - e^{-P \cdot O_j}) \quad , \forall i, j \quad (5.1)$$

$T_{ij}$  : Number of trips from  $i$  to  $j$

$E_i$  : Trip production at  $i$

$P$  : Probability of choosing a potential opportunity

$O_j$  : Number of spatially cumulative opportunities between  $i$  and  $j$ , including  $j$

$O_{j-1}$  : Number of spatially cumulative opportunities between  $i$  and  $j$ , excluding  $j$

$\alpha$  : Adjustment coefficient

A single-constrained model on trip production is supposed, presented as:

$$\sum_{l=1}^j T_{il} = E_i \quad , \forall i \quad (5.2)$$

Where  $T_{il}$  is the number of trips from  $i$  to  $l$ .

If  $T_{ij}$  in Equation 5.2 is substituted by the general form of  $T_{ij}$  in Equation 5.1, the single-constrained IOM is derived as:

$$T_{ij} = \beta \cdot E_i \cdot \frac{e^{-P \cdot O_j - 1} - e^{-P \cdot O_j}}{1 - e^{-P \cdot O_j}}, \forall i, j \quad (5.3)$$

Where  $O_j$  is the total number of opportunities and  $\beta$  is an adjustment coefficient.

Trip generation,  $E_i$  in Equation 5.3, is set as a linear function combining sociodemographic, socioeconomic, and PTLOS variables. In this study,  $E_i$  for work trips is considered as:

$$E_i^{\text{Work}} = \alpha_1^{\text{Work}} \cdot P_{1i} + \alpha_2^{\text{Work}} \cdot P_{2i} + \alpha_3^{\text{Work}} \cdot P_{3i} + \beta_1^{\text{Work}} \cdot \text{LOS}_{1i} + \beta_2^{\text{Work}} \cdot \text{LOS}_{2i} + \gamma^{\text{Work}} \cdot \text{INC}_i \quad (5.4)$$

$E_{i\text{work}}$  : Work trip production at  $i$

$P_{1i}$  : Population group age 0 to 19 years old at  $i$

$P_{2i}$  : Population group age 20 to 64 years old at  $i$

$P_{3i}$  : Population group age 65 years old and older at  $i$

$\text{LOS}_{1i}$  : Total number of PT Passage-Stops per 24 h at  $i$

$\text{LOS}_{2i}$  : Spatial density of PT stops at  $i$  (number of stops at  $i$ /area of  $i$ )

$\text{INC}_i$  : Mean income per person at  $i$

$\alpha_{1\text{work}}, \alpha_{2\text{work}}, \alpha_{3\text{work}}, \beta_{1\text{work}}, \beta_{2\text{work}}, \gamma_{\text{work}}$  : Model parameters for work trips

Trip production values for study trips are also presented with a similar equation. Afterward, a linear weighted formulation is supposed for  $T_{ij}$  as follows:

$$T_{ij} = k_W \cdot T_{ij}^{\text{Work}} + k_S \cdot T_{ij}^{\text{Study}}, \forall i, j \quad (5.5)$$

$T_{ij}$  : Total number of work and study trips

$T_{ij\text{work}}, T_{ij\text{study}}$  : Number of work and study trips from  $i$  to  $j$ , respectively

$k_W$  : Weighting coefficient for work trips

$k_S$  : Weighting coefficient for study trips

In regard to the IIOM, calibration is done with an integrated data set of work and study trips in statistical software STATA by a nonlinear optimization procedure. Results as presented in Table

5.2 show that all model parameters are statistically significant with acceptable confidence intervals.

Observed values of  $T_{ij}$  versus the estimated values can be presented as follows:

$$T_{ij} = 0,9992T_{ij}^* \quad \text{with} \quad R^2 = 83,67\% \quad (5.6)$$

Where  $T_{ij}$  is the observed number of work and study trips from  $i$  to  $j$ , and  $T_{ij}^*$  is the estimated number of work and study trips from  $i$  to  $j$ .

### 5.4.3 Modeling Performance Analysis

After presenting the coefficient of determination ( $R^2$ ) for the IIOM, performance of the model is analyzed by several Goodness-of-Fit measures. First the real and estimated cumulative values of  $T_{ij}$  versus trip duration are compared. Afterward is a discussion of performance of the IIOM in reproducing the observed values of trip production and attraction, followed by a detailed analysis of the IIOM with different Goodness-of-Fit measures.

Tableau 5.2 IIOM Calibrated Parameters

IIOM Parameters	Work	Study
<b>k</b>	0 ,573979	0 ,0933546
<b><math>\alpha_1</math></b>	-0 ,4265714	1 ,067614
<b><math>\alpha_2</math></b>	0 ,4838109	-0 ,4852591
<b><math>\alpha_3</math></b>	-0 ,075471	2 ,001871
<b><math>\beta_1</math></b>	0 ,1788247	-0 ,2829593
<b><math>\beta_2</math></b>	-41 ,75296	88 ,87629
<b><math>\gamma</math></b>	-0 ,0061696	-0 ,0721553
<b>P</b>	0 ,000151%	0 ,022300%

#### 5.4.3.1 Cumulative Number of Trips versus Trip Duration

Figure 5.4 shows the observed and estimated cumulative number of trips versus trip durations. The figure confirms that IIOM and OD values have very similar curves for trip durations that are less than about 60 min. For trip durations more than 60 min, the IIOM slightly overestimates the number of trips.

### 5.4.3.2 Trip Production and Trip Attraction Satisfaction

Evaluation of the estimated numbers for trip production ( $E_i$ ) and trip attraction ( $A_j$ ) is of great interest for understanding the appropriateness of a trip distribution model. This study calibrated the IOM with single constraints on trip production values. These values are reported from the IOM calibration:

$$E_i = 0,9684E_i^* \quad \text{with} \quad R^2 = 94,23\% \quad (5.7)$$

$$A_j = 0,9695A_j^* \quad \text{with} \quad R^2 = 92,52\% \quad (5.8)$$

$E_i$  : Observed value of trip production at  $i$

$E_i^*$  : Estimated value of trip production at  $i$

$A_j$  : Observed value of trip attraction at  $j$

$A_j^*$  : Estimated value of trip attraction at  $j$

Equations 5.7 and 5.8 confirm the appropriateness of the IOM for reproducing the trip production and attraction values.

### 5.4.4 Goodness-of-Fit Measures

In this section, several Goodness-of-Fit measures are presented that compare entries in the observed and estimated matrices. To provide an idea about the values for the IOM, also presented are Goodness-of-Fit measures of the GM calibrated with the same data set in previous studies (Nazem et al., 2012).

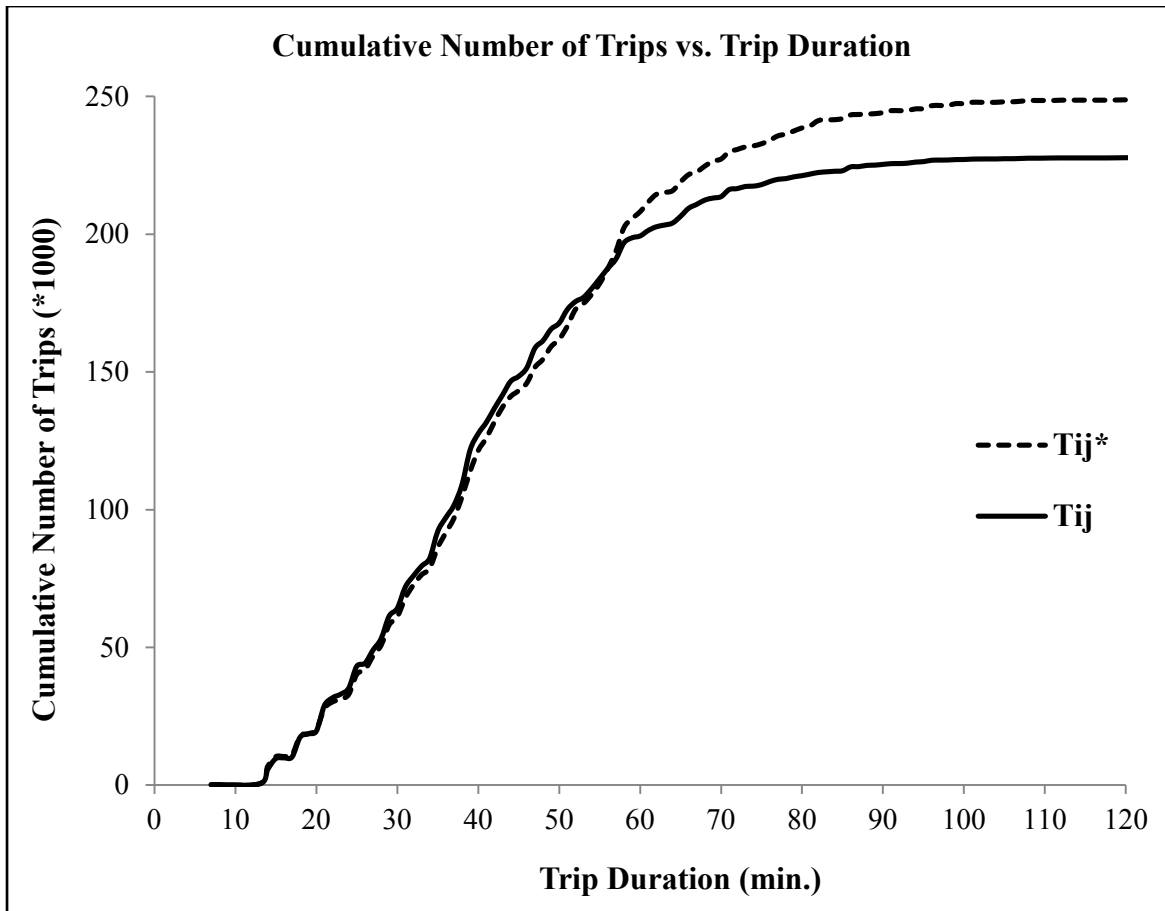


Figure 5.4 Cumulative Number of Trips versus Trip Duration

#### 5.4.4.1 Mean Trip Duration Error

Mean trip duration error is the difference between the mean duration estimated by the model and the mean trip duration based on the OD survey. In other words, trip duration for each OD is calculated from the GTFS data, and then the error is calculated by means of observed and estimated  $T_{ij}$ . The error values reported in Table 5.3 show that the IIOM is more accurate than the GM.

#### 5.4.4.2 Coefficient of Determination

Although some studies showed that in some cases the coefficient of determination or  $R^2$  may yield artificially high values in Goodness-of-Fit applications, this study presents it as a traditional measure because it is one of the most cited measures in the literature

(Black, 1991; Knudsen & Fotheringham, 1986; Smith & Hutchinson, 1981; Yaldi et al., 2011). The values of  $R^2$  reported in Table 5.3 show that the IOM is more accurate than the GM.

Tableau 5.3 Goodness-of-Fit Measures

Performance Analysis Measures		IOM	GM
Global Measures	Mean trip duration error	2,47%	4,15%
	$T_{ij}=k.T_{ij}^*$ (Value of k)	0,9992	1,0018
	$R^2$	83,67%	78,73%
	PHI	92 634,7	94 503,7
	MAE	144,58	153,32
	NMAE	0,448	0,474
	DI (PME)	23,57	23,73
Trip Production Measures	RMSE	242,9	250,2
	$E_i=k.E_i^*$ (Value of k)	0,9684	0,8996
	$R^2$	94,23%	99,67%
Trip Attraction Measures	$A_j=k.A_j^*$ (Value of k)	0,9596	1,0401
	$R^2$	92,52%	93,43%

#### 5.4.4.3 PHI Statistics

The PHI statistics are presented as follows (Sherali, Narayanan, & Sivanandan, 2003; Smith & Hutchinson, 1981):

$$PHI = \sum_{i,j} (\max\{1, T_{ij}\} \left| \ln \frac{\max\{1, T_{ij}\}}{\max\{1, T_{ij}^*\}} \right|) \quad (5.9)$$

Larger values of PHI present poorer model fits. The reported values in Table 5.3 show that the IOM, compared with the GM, yields smaller values for PHI statistics. That means it has a better performance.

#### 5.4.4.4 Mean Absolute Error and Normalized Mean Absolute Error

Mean absolute error (MAE) and normalized mean absolute error (NMAE) are defined as follows (Smith & Hutchinson, 1981):

$$MAE = \frac{\sum_{i,j} |T_{ij} - T_{ij}^*|}{N} \quad (5.10)$$

$$NMAE = \frac{MAE}{T/N} \quad (5.11)$$

$T_{ij}$  : Observed number of trips from  $i$  to  $j$ ,

$T_{ij}^*$  : Estimated number of trips from  $i$  to  $j$ ,

$T$  : Total number of trips derived from OD survey, and

$N$  : Number of estimated OD pairs.

As for PHI, larger values of the mean absolute error and normalized mean absolute error represent less accurate model fit. These measures also show that the IOM is more accurate than the GM.

#### 5.4.4.5 Dissimilarity Index or Percentage Misallocated Error

The dissimilarity index (DI) or percentage misallocated error, showing the percentage of the flows that are allocated to wrong cells in the matrix, is defined as follows (Evans & Pooler, 1987; Hu & Pooler, 2002):

$$DI = \frac{50}{T} \sum_{i,j} |T_{ij} - T_{ij}^*| \quad (5.12)$$

Larger values of the dissimilarity index show larger dissimilarities between the estimated and the observed OD survey matrices. Table 5.3 shows that on the basis of the DI the IOM behaves slightly better (i.e., has slightly better prediction capabilities) than the GM.

#### 5.4.4.6 Root Mean Square Error

The root mean square error (RMSE) is defined as follows (Yaldi et al., 2011):

$$RMSE = \sqrt{\frac{\sum_{i,j} (T_{ij} - T_{ij}^*)^2}{N}} \quad (5.13)$$



Table 5.3 confirms that the IOM behaves better than the GM, on the basis of route mean square error values.

#### **5.4.4.7 Trip Production and Trip Attraction Measures**

The coefficients of determination values presented for trip production and trip attraction in Table 5.3 show that the IOM is more efficient for reproducing trip production values. In regard to trip attraction values, both models behave in almost the same way.

On the basis of the presented Goodness-of-Fit measures, the IOM is more accurate than the classical GM. Moreover, the integrated supply-dependent formulation developed in this study allows policy makers and analysts to study potential changes in PT trip pattern because of modifications in demography, economy, and job and study spatial location, plus the PT supply.

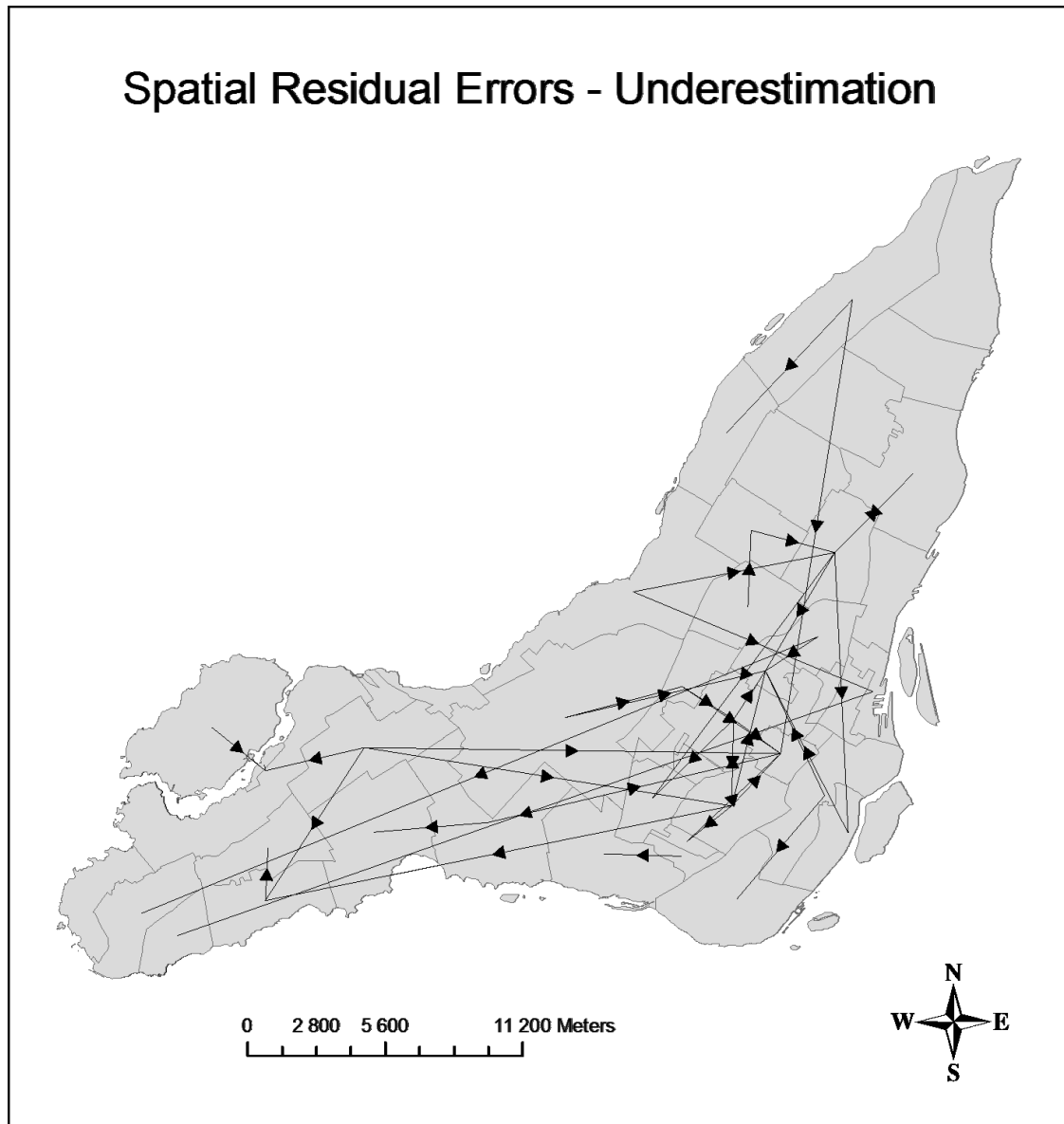
### **5.4.5 Spatial Limitations of IOM**

This section discusses limitations of the IOM from a spatial point of view to understand weaknesses and strengths of the model. Figure 5.5 shows the spatial distribution of overestimation and underestimation residual errors for the IOM by means of desire lines plotted between MS for errors between 50 % and 100 %. The following findings are highlighted.

The IOM underestimates  $T_{ij}$  more for OD pairs in which origin and destination are far away. Also seen is that most of this type of error occurs between non-central and central areas.

Most of overestimation errors reported from the IOM occur for close-distance OD pairs. One can explain the occurrence of these types of errors in IOM by the fact that the IOM family assigns trips to nearer potential opportunities. A possible solution to prevent this type of error might be the development of a more accurate method to define intervening opportunities for each trip purpose, for in this study, there was use of only PT trip duration for calculating the number of intervening opportunities.

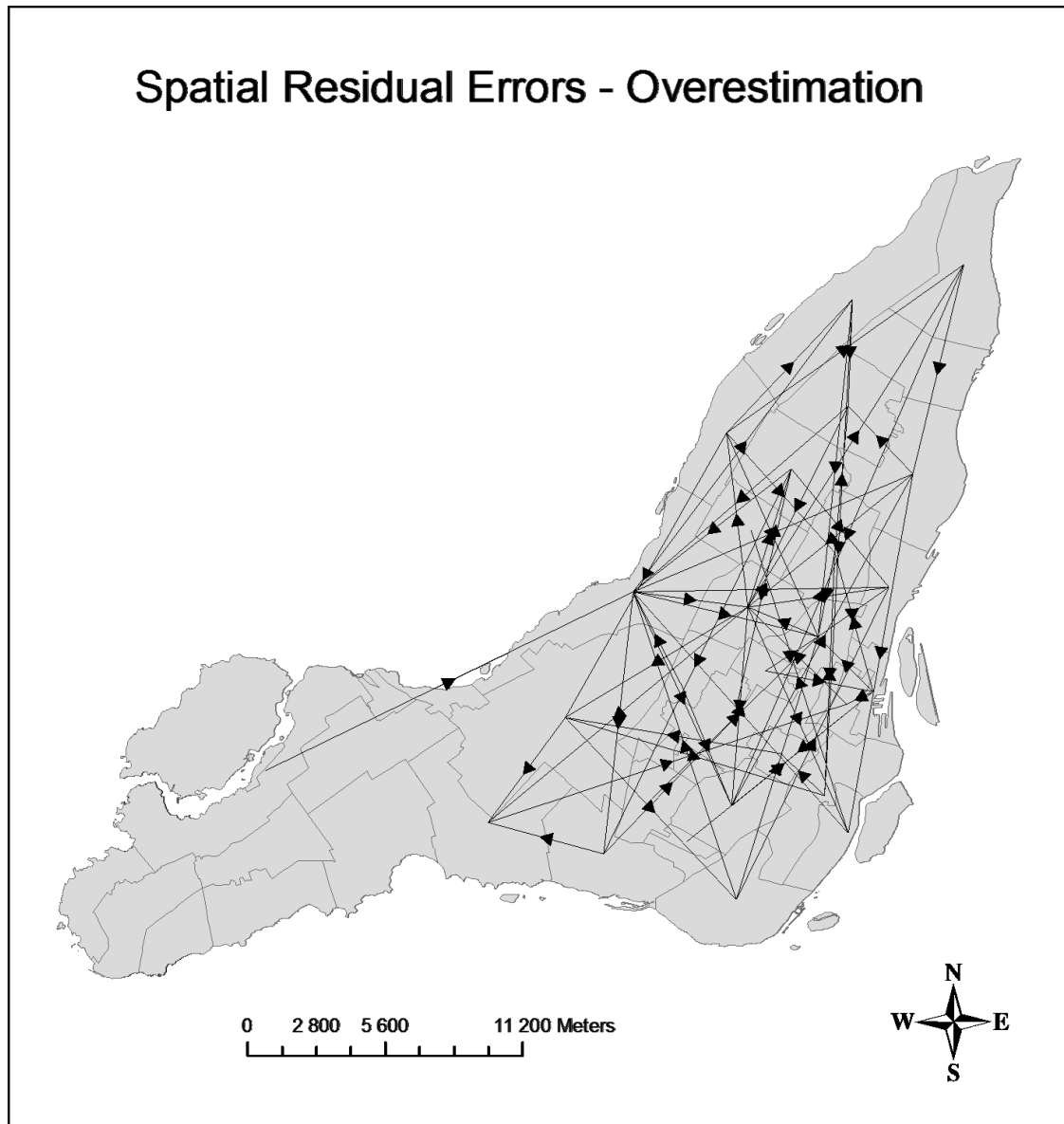
This study also performed the same spatial residual error analysis for the GM calibrated as in previous studies (Nazem et al., 2012). Random underestimation and overestimation errors were found for both far-and-close-distance OD pairs.



(a)

Figure 5.5 Spatial Residual Errors Reported from IIOM

(a) Underestimation; and (b) Overestimation



(b)

Figure 5.5 Spatial Residual Errors Reported from IIOM (suite)

(a) Underestimation; and (b) Overestimation

## 5.5 Conclusion and Future Research

In this study, an integrated generation–distribution model that is supply-dependent was developed for PT trips. This new model, called the IIOM, is calibrated for study and work purpose trips with single constraints on trip production values for PT morning peak hour's trips in the Island of Montreal. Analysis of modeling performance by means of several Goodness-of-Fit measures showed that the IIOM is well behaved (i.e., globally it has good prediction capabilities), and it is more accurate than the classic GM presented in previous studies (Nazem et al., 2012).

Also studied were limitations of the IIOM from a spatial point of view. In other words, by comparing observed and estimated trip numbers on an underestimation-overestimation basis, spatial weaknesses of the IIOM are known. This could help in finding improvement strategies in regard to explanatory variable choice or even model formulation modifications; those areas are part of the authors' ongoing research.

The new formulation of the IIOM represents great potential for policy makers or PT analysts, owing to the sociodemographic, socioeconomic, and PT supply variables.

As all the required data for calibrating the IIOM come from external and independent sources, the IIOM could be used for the sake of PT trip generation and distribution forecasting for the future. More precisely, for calibrating the IIOM, one needs reference OD matrices, number of potential work and study opportunities, OD trip durations, sociodemographic, socioeconomic, and PT supply characteristics at the MS level.

The presence of sociodemographic, socioeconomic, and especially PT supply characteristics in the IIOM allows study of the effect of potential changes of these variables on the PT trip pattern. As discussed in previous studies, different sociodemographic groups have different behaviors in a PT network (Nazem et al., 2011). In regard to the PT supply characteristics, the developed IIOM could be used as a tool to forecast the PT trip pattern owing to modifications in the PT network. Because in the IIOM, two Level-Of-Service variables of total number of PT Passage-Stops per 24 h at MS  $i$  and spatial density of PT stops at MS  $i$  are used, the analyst could verify the effect of different types of PT network modifications on the PT trip pattern.

In regard to required trip duration values for calibrating trip distribution models, it is supposed intuitively that the IIOM compared with the GM needs less accurate data. In fact, in the IIOM, trip duration is used only to rank zones according to distance, but in the GM, trip duration values are used explicitly in the model formulation. In contrast, studying the sensitivity of the IIOM to

trip duration values is of great interest because a minor change in the trip duration can change the ranking of zones and that will result in changes in the number of intervening opportunities.

For further research, a bi-level optimization approach is proposed by using smartcard data. At the first level, MS trip production ( $E_i$ ) will be defined based on sociodemographic, socioeconomic, and PTLOS variables. Then it will be calibrated with data derived from smartcard validations. The second level is dedicated to the calibration of an IOM by means of the OD survey data and trip production values calculated at the first level. This new model could be interesting, because it allows the use the additional data derived from smartcard validations in the GMA.

Another research topic is development of the IIOM at a grid level. Because the present study is done at the MS level, a grid study could present interesting results to compare the sensitivity of PT trip distribution models with the study level.

It is concluded that the use of the newly developed IIOM is advantageous owing to its behavioral and supply-dependent bases. This paper presents the first step in using the IIOM for PT trip analysis, planning, and forecasting in the GMA. Evidently, it will be followed in the future by the development of more accurate models by means of continuous smartcard data.

## Acknowledgments

The authors acknowledge the support of Natural Sciences and Engineering Research Council of Canada and of the Agence Métropolitaine de Transport. They thank the survey consortium that provided the Greater Montreal Origin–Destination Survey and the GTFS data. The authors especially thank François Godefroy, Audrey Godin, Pierre-Léo Mongeon-Bourbonnais, Félix Pépin, Éric Poliquin, Hubert Verreault, and other members of the Mobilité Chair of École Polytechnique de Montréal for their help in obtaining and preparing the required data sets.

# CHAPITRE 6 : ARTICLE 3 : « REVISITING THE DESTINATION RANKING PROCEDURE IN DEVELOPMENT OF AN INTERVENING OPPORTUNITIES MODEL FOR PUBLIC TRANSIT TRIP DISTRIBUTION »

rédigé par :

Mohsen Nazem<sup>7</sup>, Martin Trépanier<sup>8</sup>, et Catherine Morency<sup>9</sup>

soumis pour publication dans :

Journal of Geographical Systems

date de soumission initiale : 12 juillet 2013

date de recevoir des commentaires de comité de lecture : 24 décembre 2013

date de soumission de la version révisée : 15 mai 2014

## Abstract

An Enhanced Intervening Opportunities Model (EIOM) is developed for Public Transit (PT) trips. This is a distribution supply-dependent model, with single constraints on trip production for work trips during morning peak hours (6:00 AM to 9:00 AM) within the Island of Montreal, Canada. Different data sets including the 2008 Origin-Destination (OD) survey of the Greater Montreal Area, 2006 Census of Canada, GTFS network data, along with the geographical data of the study area are used. EIOM is a nonlinear model composed of sociodemographic, PT supply and work location attributes. An enhanced destination ranking procedure is used for calculating the number of spatially cumulative opportunities, the basic variable of EIOM. For comparison, a Basic Intervening Opportunities Model (BIOM) is developed by using the basic destination

---

7 M. Nazem (Corresponding Author), Ph.D. Candidate. Department of Mathematical and Industrial Engineering, École Polytechnique de Montréal. C.P. 6079, Station Centre-Ville, Montreal, Quebec H3C 3A7, Canada. (mohsen.nazem@polymtl.ca)

8 M. Trépanier, Professor. Department of Mathematical and Industrial Engineering, École Polytechnique de Montréal. C.P. 6079, Station Centre-Ville, Montreal, Quebec H3C 3A7, Canada. (mtrepanier@polymtl.ca)

9 C. Morency, Associate Professor. Department of Civil, Geological, and Mining Engineering, École Polytechnique de Montréal. C.P. 6079, Station Centre-Ville, Montreal, Quebec H3C 3A7, Canada. (cmorency@polymtl.ca)

ranking procedure. In fact, the main difference between EIOM and BIOM is in the destination ranking procedure; EIOM considers the maximization of a utility function composed of PT Level-Of-Serve (LOS) and number of opportunities at the destination, along with the OD trip duration, whereas BIOM is based on a destination ranking derived only from OD trip durations. Analysis confirmed that EIOM is more accurate than BIOM. This study presents a new tool for PT analysts, planners and policy-makers to study the potential changes in PT trip patterns, due to changes in sociodemographic characteristics, PT supply, etc. Also it opens new opportunities for the development of more accurate PT demand models with new emergent data such as smart card validations.

## **6.1 Introduction**

Trip distribution is the second step in the classical sequential four step models for aggregate transport planning. Several families of trip distribution models such as Gravity Model (GM) and Intervening Opportunities Model (IOM) are presented in the literature (Bonnell, 2004; Hensher & Button, 2008; Ortuzar & Willumsen, 1994; Wilson, 1970). In GM, the number of trips between each Origin-Destination (OD) pair is based directly on the OD distance (or trip duration) and trip production-attraction values, whereas the IOM considers the number of intervening opportunities as the main influencing factor. As it seems that both these factors influence the trip distribution, unified hybrid models are also developed and calibrated (Almeida & Goncalves, 2001; Goncalves & Ulysséa Neto, 1993; Wills, 1986).

Based on past studies, each family of distribution models has its own advantages and disadvantages compared to other models. At the same time, the calibration of each model requires different data sets and calibration procedures. The IOM has been used less than the GM due to complexities in terms of the calibration procedure (Goncalves & De Cursi, 2001). However compared to the GM, the IOM is behavioral based (Kermanshah, 2004), less sensitive to the size and shape of the study area (Chow et al., 2005) and also produces better results in cases where destinations which satisfy the trip purpose are not uniformly distributed (Veenstra et al., 2010). In some previous studies, the IOM was calibrated with several different approaches for the Island of Montreal, Canada, and the results showed that IOM has better performance than GM, in terms of trip production and trip attraction and other Goodness-of-Fit measures (Nazem et al., 2012, 2013).

For the calibration of most distribution models, an OD reference matrix is required. Several approaches are presented in the literature, such as asking passengers to fill out questionnaires on board of transit vehicles, estimating the number of passengers from counting the boarding and alighting passengers at stops, or from smart card validation data (S. Feng & Li, 2004). In this study, we use an OD matrix obtained from 2008 OD survey held in the Greater Montreal Area (GMA); this survey reaches some 5 % of the residing households.

The study area that is presented in Figure 6.1 consists of 41 Transportation Analysis Zones (TAZ) which are Montreal Municipal Sectors (MS) with a population of nearly 1,9 millions, distributed over almost 500 km<sup>2</sup> (StatCan, 2011b). The present study is limited to the Island of Montreal and Public Transit (PT) work trips in order to deal with less calculation complexities, remembering that the methodology can be generalized to the whole GMA.

As for the IOM, the number of intervening opportunities represents the most influencing factor, the procedure of its calculation becomes essential. For the calibration of an IOM, from each origin zone we need to rank all destination zones based on their relative attractiveness. Until now, most studies consider trip time or distance for representing the attractiveness of each destination zone (Nazem et al., 2012, 2013). But it seems that trip duration or distance is not the only influencing factor for public transit trips; other factors like PT Level-Of-Service (PTLOS) and number of potential opportunities at the destination zone could also be relevant. Thus, we propose a new ranking procedure called “enhanced destination ranking procedure” based on OD trip duration, PTLOS and number of potential opportunities at the destination zone.

In this study, we aim to calibrate an Enhanced Intervening Opportunities Model (EIOM) with the enhanced destination ranking procedure, based on the economic consumer theory. This model has great potentials for practitioners especially in fields like “seeding” population and work places due to new major transportation projects and urban development. Another potential of EIOM is related to its supply-dependence nature. EIOM could be used as a tool for directly evaluation the PT travel pattern due to changes in PTLOS. The LOS independent variable that is derived from the GTFS data could be used for representing the PT service modifications of different types, such as enhancing the actual PT service or adding new PT services. Also the population by age cohort variables in EIOM could be used for studying the effects of sociodemographic changes on PT trip pattern in the study area. The latter possibility could be interesting especially in developed countries because of population aging issues.



The main advantage of EIOM compared to BIOM is that the PTLOS and population by age cohort independent variables are used explicitly in the model formulation, and could be modified as model inputs. In classical models these variables are used more implicitly as variables that cause changes in other variables.

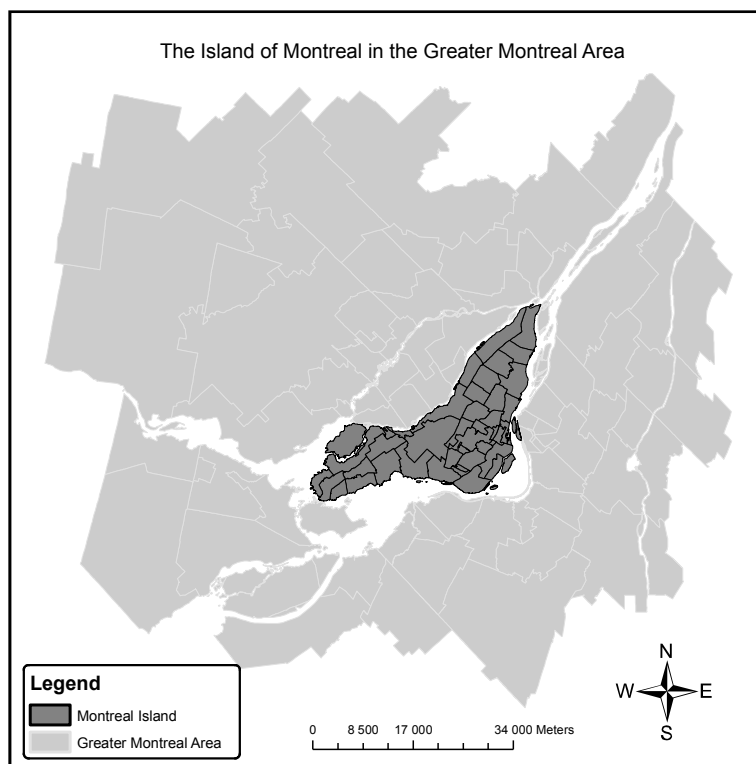


Figure 6.1 Spatial Distribution of MS in the Island of Montreal

As in this study many acronyms are used, a list is presented here:

- BIOM: Basic Intervening Opportunities Model
- DI: Dissimilarity Index
- EIOM: Enhanced Intervening Opportunities Model
- GM: Gravity Model
- GMA: Greater Montreal Area
- GTFS: General Transit Feed Specification
- IOM: Intervening Opportunities Model
- LOS: Level-Of-Service

- MAE: Mean Absolute Error
- MS: Municipal Sector
- NMAE: Normalized Mean Absolute Error
- OD: Origin-Destination
- PME: Percentage Misallocated Error
- PT: Public Transit
- RMSE: Root Mean Square Error
- STM: Société de Transport de Montréal
- TAZ: Transportation Analysis Zone

The structure of the paper is as follows. First we present a literature review to learn more about advances in trip distribution models; then the data sets and their preparation for modeling development are described. In the next section, following a descriptive analysis of the data, the development of BIOM and EIOM is presented. Afterwards, the performance of EIOM is studied and compared to BIOM, by means of several global and Goodness-of-Fit measures. The next section presents spatial limitations of the models by means of a spatial residual errors analysis. In the conclusion, we discuss some interesting potentials of EIOM for analysts and policy-makers, and also present some of our ongoing and future research topics.

## 6.2 Literature Review

Several families of models are used for estimating trip distribution in its general aggregate form. Among them, the Gravity Model (GM) and the Intervening Opportunities Model (IOM) are the most common (Bonnell, 2004; Ortuzar & Willumsen, 1994).

The GM for trip distribution inspired by Newton's law uses an impedance function, which is generally represented by a generalized cost. This model and its applications are well presented in the literature (Bonnell, 2004; De Grange et al., 2009; Hensher & Button, 2008; Ortuzar & Willumsen, 1994; Rajesson, 2009; Thamizh Arasan et al., 1996; Trépanier, Morency, & Rajesson, 2008; Wilson, 1970).

The main idea of the opportunity model came from some theoretical concepts that relate the mobility, the migration distances and the spatial locations of services; the theory of this model in its present form was developed later (Schneider, 1959; Stouffer, 1940, 1960). The fundamental

idea of this model is that distance is not the main factor that affects destination choices. Contrariwise, this model considers the main influencing factor as the relative accessibility of the opportunities that can satisfy each trip purpose. This model assumes that an individual chooses the closest destination location that gives him the opportunity to meet his needs. Distance or more widely, generalized cost, is not a continuous variable as it was used explicitly in gravity model, and it serves rather to find the ranking of destinations from a given origin point (Bonnell, 2004; Hensher, 1977; Hensher & Button, 2008; Ortuzar & Willumsen, 1994). The use of this model in transportation planning is briefly presented in literature. In the 1980s, opportunity model was used for modeling during the Chicago Area Transportation Studies (Eash, 1983, 1984). More recently, the IOM was used for simulating student flows and results confirmed such model has better performance than the GM (Almeida & Goncalves, 2001). In order to consider both distance and intervening opportunities in a single trip distribution model, a hybrid gravity-intervening opportunities model is also presented and tested in the literature (Goncalves & Ulysséa Neto, 1993; Wills, 1986). Also a recent study shows that a destination choice model based on the utility maximization principle behaves better than the GM for reproducing an observed OD matrix (Mishra, Wang, Zhu, Moeckel, & Mahaparta, 2013). An Intergraded IOM (IIOM) which integrates trip generation and trip distribution in a supply-dependant model is calibrated on PT trips within the Island of Montreal, and the results showed that the IIOM behaves better than the GM in order to reproduce an observed OD matrix (Nazem et al., 2013). In the current study we aim to develop and calibrate an EIOM, based on an enhanced destination ranking procedure, and compare its performance to BIOM which used the basic destination ranking procedure, based only on PT trip durations.

For calibrating trip distribution models, reference OD matrices are generally required, which could be estimated by several methods; direct observation, synthesis and etc. (Yaldi et al., 2011). In this study, we obtain reference OD matrices by processing the data collected during a large-scale OD survey held in 2008 in the GMA (AMT, 2010).

After calibration of trip distribution models, we need some global and Goodness-of-Fit measures for analyzing the model performances. Several measures are presented in the literature (Akwawua & Pooler, 2001; Black, 1991; Chow et al., 2005; Evans & Pooler, 1987; Hu & Pooler, 2002; Smith & Hutchinson, 1981; Yaldi et al., 2011). We will present formulations of required measures and other related works in the following sections.

## 6.3 Data for Analysis

In this section we introduce the data sets used in this study. First we present the reference data sets and afterward, their preparation for the modeling. Figure 6.2 shows a data flow diagram of all reference and prepared data sets.

### 6.3.1 Reference Data Sets

Reference data sets represent the raw data used for modeling. In this study, we used the 2008 GMA OD survey, the 2006 census of Canada, the GTFS files, along with the geographical data of the study area at the MS level.

#### 6.3.1.1 GMA Origin-Destination Survey

For almost forty years, the GMA authorities have been conducting telephone OD travel surveys approximately every five years. This data includes rich information regarding all trips made by every person in a 5 % sample of residing households, which makes the OD survey “the primary source of information on people movement habits” (AMT, 2010). Precise spatiotemporal details are collected on all-purpose and all-mode trips. In this study, we used the data coming from the 2008 OD survey. The sample contains almost 319 900 trips. Sociodemographic information such as dwelling location, household size, car ownership and class of income and age, gender and main occupation are also gathered. For each record an expansion factor is calculated that will be used to expand the data set to all population based on the collected 5 % sample (AMT, 2010). The OD survey is used for deriving the reference OD matrices for work trips, and also for the sake of trip production/attraction performance analysis.

#### 6.3.1.2 Census of Canada

Census of Canada is a “unique undertaking on a vast scale” conducted every five years by Statistics Canada (StatCan, 2011b). It consists of collecting data from almost 31,6 million people and more than 13,5 million dwellings. Since the travel data is calibrated by using the corresponding population, we used the data from the 2006 census, for deriving the population per age group and also the number of opportunities for work trips at the MS level (StatCan, 2011a).

### 6.3.1.3 General Transit Feed Specifications (GTFS)

Based on the definition given by Google, “the GTFS defines a common format for PT schedules and associated geographic information. GTFS allow PT agencies to publish their transit data and developers to write applications that consume that data in an interoperable way” (Google, 2012). In this study we used the GTFS data of the study area that is obtained from Société de Transport de Montréal (STM), the PT agency of the Island of Montreal.

The GTFS is used for several purposes such as characterizing the PT Level-Of-Service in each MS, and also calculating the PT trip duration for each OD pair. The PTLOS is represented with the total number of Passage-Stops (transit vehicle passing a stop) per 24 hours in each MS, which is a variable sensitive to changes in PT supply. As the GTFS data is more and more publicly available, this could be a new approach for measuring PTLOS in an aggregate way by using the detailed GTFS data. Some studies recently used the GTFS data for measuring PT performance (Wong, 2013).

For calculating PT trip durations for each OD pair, as this study is aggregate at MS level, we consider the geographic centroid of each MS as its spatial delegate. Google directions API and morning peak hour were used to get the shortest routes and related trip durations including walking access and egress, waiting, transfer and in-vehicle times. As the PTLOS remains almost constant during the morning period within the Island of Montreal, the PT trip duration is calculated at 8:00 AM. Based on this hypothesis, we obtained a PT trip duration matrix between all 41 MS in the Island of Montreal. The next section presents the data preparation.

## 6.3.2 Data Preparation

In this section, we discuss the data preparation for calibrating the models. Based on the presented data sources in Figure 6.2, we obtain a reference OD matrix, number of opportunities in each MS, number of spatially cumulative opportunities for each OD pair for BIOM and EIOM, and the total number of Passage-Stops per 24 hours in each MS.

### 6.3.2.1 Reference OD Matrix

The OD matrix for work trips is derived from the 2008 OD survey data. Each record in the OD survey presents complete characteristics of a person’s trips. First we exclude all return-home and off-peak trips to obtain the required data. Afterwards the number of work trips for each OD pair

is calculated by summing up the expansion factors for all concerned OD records in the survey. As the OD survey is limited to a 5 % household sample, some OD pairs will have no trips, meaning that there is no data in the survey for those OD pairs.

This kind of OD reference data set is usually presented in a matrix form, but in this study we turned columns into rows to obtain a table form (Sample presented in Table 6.2). Afterwards, trip production and attraction for each MS are obtained for analyzing the modeling performance.

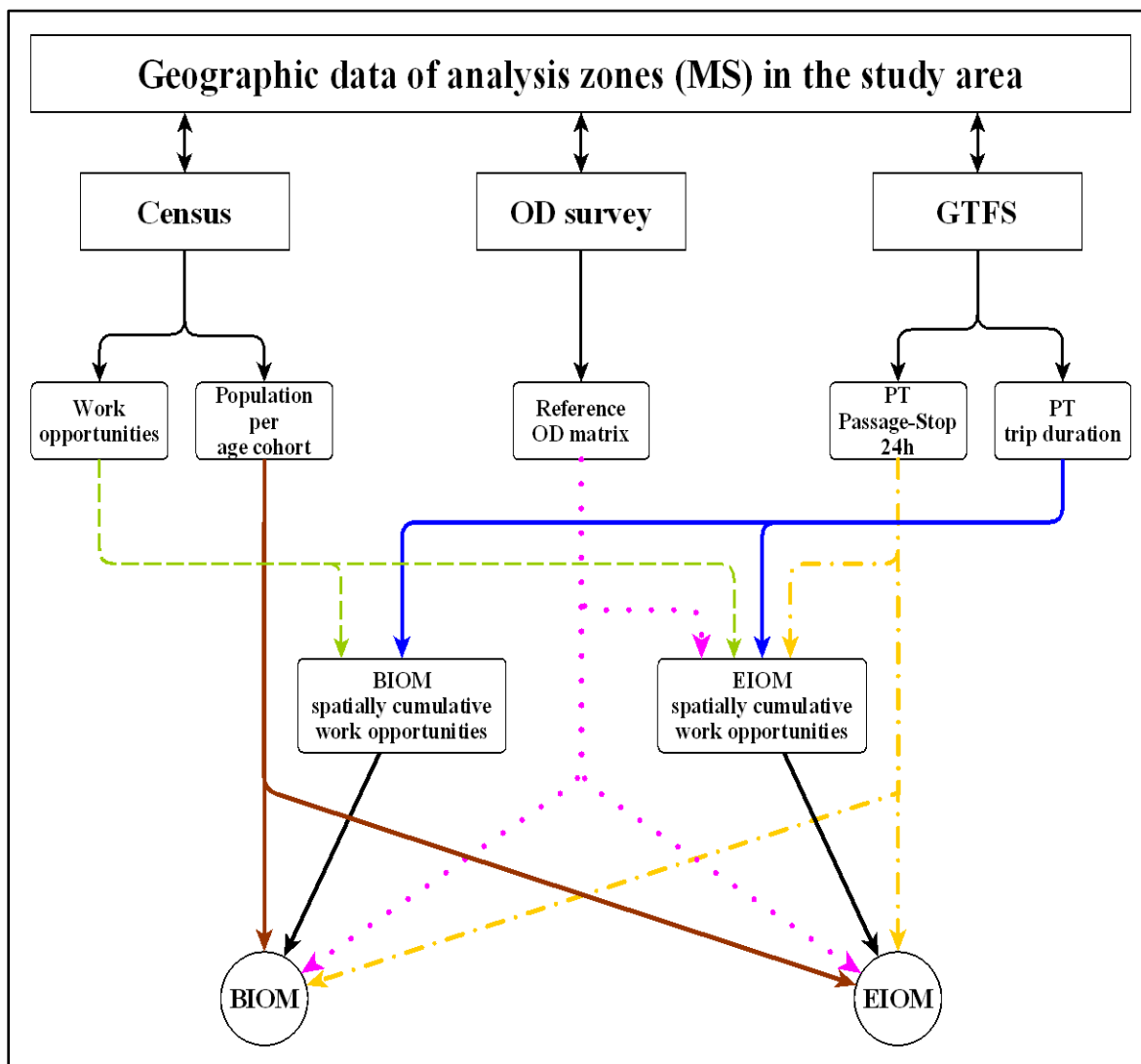


Figure 6.2 Data Flow Diagram for Development of BIOM and EIOM

### 6.3.2.2 Number of Opportunities in Each MS

The development of trip distribution models needs the number of opportunities in each MS. As number of opportunities that can satisfy each trip purpose depends on activity locations that can fulfill that activity type, we need a data set that presents the number of work opportunities at the MS level. This is derived from the 2006 census of Canada.

### 6.3.2.3 Number of Spatially Cumulative Opportunities in Each MS

For the development of an IOM, we need the number of intervening opportunities for each OD pair. The number of intervening opportunities between zones  $i$  and  $j$  is the sum of all opportunities located *between*  $i$  and  $j$ . This number is called spatially cumulative opportunities, and will be used in the calibration of the IOM (Equation 6.5). In this study, the number of spatially cumulative opportunities is calculated in two different ways, which will result in two different models called BIOM and EIOM:

- BIOM: The calculation of spatially cumulative opportunities is done by the basic destination ranking procedure based only on PT trip durations.
- EIOM: The calculation of spatially cumulative opportunities is done by the enhanced destination ranking procedure based on the maximization of a utility function composed of PT trip duration, PTLOS and number of potential opportunities at each destination. In contrast with BIOM where only trip duration was used for ranking all possible destinations from each given origin, in EIOM, values of the utility function for each OD pair are used for ranking.

Figure 6.2 shows a data flow diagram for the development of both BIOM and EIOM. In next section, the development of the models is described.

## 6.4 Development of the Models

First the general form of the IOM is presented, followed by calibrating BIOM and EIOM and analysis of their performance.

Both models will be calibrated and validated for PT work trips during average weekday morning peak hours (6:00 AM - 9:00 AM) within the Island of Montreal. After applying the expansion

factors in the 2008 GMA OD survey, we obtain almost 158 790 PT work trips, which represent 57,1 % of all PT morning peak trips in the study area.

In EIOM, we aim to calculate the coefficients of the utility function of the enhanced destination ranking procedure by means of a conditional Logit model for alternative specific variables. Afterwards based on this utility function, we rank all destination zones from each origin zone and calculate the number of intervening opportunities for each OD pair. EIOM is calibrated based on the new number of intervening opportunities.

The advantage of EIOM is that it allows policy-makers to study the effects of PT supply modifications on trip pattern. Also as the required input data for this model could be available for the future, it could be used for forecasting purposes.

### 6.4.1 General Form of IOM

The IOM in its general form is presented as (Bonnell, 2004):

$$T_{ij} = k \cdot E_i \cdot (e^{-P \cdot O_{j-1}} - e^{-P \cdot O_j}) \quad , \forall i, j \quad (6.1)$$

$T_{ij}$  : Number of trips from  $i$  to  $j$

$E_i$  : Trip production at  $i$

$P$  : Probability of choosing a potential opportunity

$O_j$  : Number of spatially cumulative opportunities between  $i$  and  $j$ , including  $j$

$O_{j-1}$  : Number of spatially cumulative opportunities between  $i$  and  $j$ , excluding  $j$

$k$  : Adjustment coefficient

We suppose a single-constrained model on trip production, presented as:

$$\sum_{l=1}^J T_{il} = E_i \quad , \forall i \quad (6.2)$$

If we substitute the  $T_{ij}$  in Equation 6.2 with the general form of  $T_{ij}$  in Equation 6.1, the single-constrained IOM is derived as:



$$T_{ij} = k \cdot E_i \cdot \frac{e^{-P \cdot O_{j-1}} - e^{-P \cdot O_j}}{1 - e^{-P \cdot O_j}}, \forall i, j \quad (6.3)$$

Where  $O_j$  is total number of opportunities.

Trip generation,  $E_i$  in Equation 6.3 is set as a linear function combining sociodemographic and PTLOS characteristics. In this study, we consider  $E_i$  as:

$$E_i = a_1 \cdot P_{1i} + a_2 \cdot P_{2i} + a_3 \cdot P_{3i} + b \cdot LOS_i, \forall i \quad (6.4)$$

$P_{1i}$ : Population age between 0-19 years old at  $i$

$P_{2i}$ : Population age between 20-64 years old at  $i$

$P_{3i}$ : Population age 65 years old and older at  $i$

$LOS_i$ : Total number of PT Passage-Stops per 24 hours at  $i$

$a_1, a_2, a_3, b$ : Model parameters

By combining Equations 6.3 and 6.4, the general form of the IOM is derived as:

$$T_{ij} = k \cdot (a_1 \cdot P_{1i} + a_2 \cdot P_{2i} + a_3 \cdot P_{3i} + b \cdot LOS_i) \cdot \frac{e^{-P \cdot O_{j-1}} - e^{-P \cdot O_j}}{1 - e^{-P \cdot O_j}}, \forall i, j \quad (6.5)$$

Both BIOM and EIOM have the same form (as presented in Equation 6.5); the only difference between them is the way that the destination ranking procedure is applied. Supposing  $k=1$ , in next sections we describe the development of BIOM and EIOM for work trips.

### 6.4.2 BIOM

BIOM is based on the formulation presented in Equation 6.5. As the model in this study is calibrated on work trips, we need the number of work opportunities in each MS, which is derived from the 2006 census of Canada. Regarding the intervening opportunities which are presented in Equation 6.5 by the number of spatially cumulative opportunities, the basic destination ranking procedure is applied.

#### 6.4.2.1 Basic Destination Ranking Procedure

After calculating PT trip durations for all OD pairs, we suppose that the number of intervening opportunities for each OD pair is the sum of all opportunities located spatially between the origin and the destination. The term “spatially” in this case is related to PT trip duration and not to the geographical position of origin and destination MS. This number could be calculated by executing the following steps:

- Step 1: Sum up all opportunities in each destination MS in order to find the total number of opportunities in each MS.
- Step 2: From each origin MS, rank all destination MS based on the increasing PT trip duration.
- Step 3: Calculate the number of spatially cumulative opportunities including each destination MS, based on the ranked MS in step 2.
- Step 4: Calculate the number of spatially cumulative opportunities excluding each destination MS, based on the ranked MS in step 2.

Table 6.1 shows the form of the derived data set for the origin MS 101.

By integrating the spatially cumulative opportunities data, OD trip matrix, and PTLOS and sociodemographic data, we derive all data that is required for calibrating BIOM. Afterwards the calibration is done by the statistical software, STATA, via a nonlinear optimization procedure; results presented in Table 6.6.

Observed values of  $T_{ij}$  versus the estimated values could be presented as:

$$T_{ij} = 0,8662T_{ij}^*, \forall i, j \quad \text{with} \quad R^2 = 84,84\% \quad (6.6)$$

$T_{ij}$  : Observed number of work trips from  $i$  to  $j$

$T_{ij}^*$  : Estimated number of work trips from  $i$  to  $j$

Tableau 6.1 Sample of Spatially Cumulative Opportunities for Work Trips (Basic Destination Ranking Procedure)

Origin	Destination	OD Index	PT Trip Duration in Ascending Order (Ranking)	Number of Opportunities at Each Destination	Number of Spatially Cumulative Opportunities between i and j, Including j	Number of Spatially Cumulative Opportunities between i and j, Excluding j
<i>i</i>	<i>j</i>	<i>ij</i>	<i>d<sub>i j</sub></i>	<i>opp<sub>j</sub></i>	<i>o<sub>j</sub></i>	<i>o<sub>j-1</sub></i>
...						
101	101	101101	0	182 215	182 215	0
	102	101102	13	106 045	288 260	182 215
	106	101106	15	43 994	332 254	288 260
	107	101107	19	48 878	381 132	332 254
	105	101105	26	65 779	446 911	381 132
	...					
	138	101138	74	3 165	1 095 350	1 092 185
	137	101137	75	9 780	1 105 130	1 095 350
	140	101140	85	4 485	1 109 615	1 105 130
	134	101134	96	1 835	1 111 450	1 109 615
	141	101141	102	1 165	1 112 615	1 111 450
...						

### 6.4.3 EIOM

For calibrating EIOM (Equation 6.5), we use the same data set for sociodemographic and PTLOS attributes; the only difference between EIOM and BIOM is the destination ranking procedure. This will affect the number of spatially cumulative opportunities.

#### 6.4.3.1 Enhanced Destination Ranking Procedure

From each origin zone we aim to rank all possible destinations based on their attractiveness. The attractiveness of each destination,  $j$ , from each origin,  $i$ , is represented by a utility function,  $U_{ij}$ , composed of several alternative specific variables. Different variables sets are examined and the most performance combination among examined variable sets is presented by:

$$U_i^j = u_1 \cdot d_{ij} + u_2 \cdot LOS_j + u_3 \cdot OPP_j, \forall i, j \quad (6.7)$$

$U_i^j$  : Utility of choosing destination zone  $j$  from origin zone  $i$

$d_{ij}$  : PT trip duration between  $i$  and  $j$  (min.)

$LOS_j$  : Total number of PT Passage-Stops per 24 hours at  $j$

$OPP_j$  : Total number of work opportunities at  $j$

$u_1, u_2, u_3$  : Model parameters

The variable LOS represents the overall PTLOS at each MS. We used this variable for destination MS in order to imply the importance of each destination in the choice set. After examining several variables such as density of Stops, etc. and several time windows, we obtained the actual definition of the PTLOS variable that is the total number of Passage-Stops per 24h for the whole MS. This definition could yield better results for calibrating the utility function in Equation 6.7.

As the study area is composed of 41 MS, we can suppose 41 different alternative destination zones from each given origin zone. Based on this hypothesis, we make a choice set of 41 MS for each declared OD trip in the 2008 OD survey. Table 6.2 shows a sample of data derived from the 2008 GMA OD survey. For creating the choice set, each line of Table 6.2 will be expanded to 41 lines, each one representing one of the 41 possible alternative destination MS in the study area. Table 6.3 shows a sample of the created choice set. In this table, the column *altj* shows the 41 alternative MS destinations, and *choice* column shows the chosen alternative. The values in the *t<sub>ij</sub>\_work* column represent the frequency of each choice in the data set, and the last three columns represent the values of the alternative specific variables in Equation 6.7.

We calibrated a conditional Logit model by the statistical software, STATA on the data set in Table 6.3 to calculate the coefficients of the utility function presented in Equation 6.7. Results that are presented in Table 6.4 show that all variables are statistically significant with 95 % confidence intervals.

Now we can calculate the utility value for each OD pair and also the probability of choosing each possible alternative destination from each given origin:

$$P_i^j = \frac{e^{U_i^j}}{\sum_{k=101}^{141} e^{U_i^k}}, \forall i, j \quad (6.8)$$

$P_i^j$ : Probability of choosing destination zone  $j$  from origin zone  $i$

$U_i^j$ : Utility of choosing destination zone  $j$  from origin zone  $i$

The enhanced ranking procedure is based on the probabilities calculated by Equation 6.8. Table 6.5 shows a sample of the new destination ranks from the origin zone 102. The fact that the trip duration values, column  $d_{i_j}$  are not anymore in the ascending order confirms the difference between the basic and the enhanced ranking procedures.

After calculating the probabilities of choosing each destination zone from each origin zone, we perform the following steps in order to calculate the number of spatially cumulative opportunities, based on the enhanced ranking procedure:

- Step 1: Sum up all opportunities in each destination MS in order to find the total number of opportunities in each MS.
- Step 2: From each origin MS, rank all destination MS based on the increasing probability calculated from Equation 6.8 (Sample presented in Table 6.5).
- Step 3: Calculate the number of spatially cumulative opportunities including each destination MS, based on the ranked MS in step 2.
- Step 4: Calculate the number of spatially cumulative opportunities excluding each destination MS, based on the ranked MS in step 2.

This will result in a data set similar to Table 6.1, but derived from the enhanced ranking procedure. By integrating this data set with the OD trip matrix, PTLOS and sociodemographic data we derive all required data for calibrating EIOM, remembering that EIOM is also based on the formulation presented in Equation 6.5. The calibration is done by the statistical software, STATA via a nonlinear optimization procedure. Results show that all variables are statistically significant with acceptable confidence intervals (Table 6.6).

Observed values of  $T_{ij}$  versus the estimated values could be presented as follows:

$$T_{ij} = 0,8923T_{ij}^*, \forall i, j \text{ with } R^2 = 88,47\% \quad (6.9)$$

In the next section, we compare the performance of BIOM and EIOM, in order to confirm the appropriateness of the enhanced destination ranking procedure.

Tableau 6.2 Sample of OD Trip Numbers Derived from the 2008 GMA OD Survey

Index Used for Modeling	Origin	Destination	Work OD Trips
<i>n</i>	<i>i</i>	<i>j</i>	<i>t<sub>ij</sub>_work</i>
1	101	101	388
2	101	102	102
3	101	106	81
4	101	107	165
5	101	105	49
...	...	...	...

Tableau 6.3 Sample of Data Layout for Conditional Logit Model (Enhanced Destination Ranking Procedure)

Index Used for Modeling	Origin	List of 41 Alternative Destinations	Chosen Alternative	Frequency of Each Choice	PT Trip Duration between Origin and Each Alternative Destination	PT Pass-Stop 24 Hours at Each Alternative Destination	Number of Opportunities at Each Alternative Destination
<i>n</i>	<i>i</i>	<i>altj</i>	<i>choice</i>	<i>t<sub>ij</sub>_work</i>	<i>d<sub>i</sub>_altj</i>	<i>los<sub>altj</sub></i>	<i>opp<sub>altj</sub></i>
...							
3	101	101	0	81	0	28 802	182 215
		102	0	81	13	24 003	106 045
		103	0	81	30	23 859	22 815
		104	0	81	27	30 880	19 892
		105	0	81	26	42 064	65 779
		106	1	81	15	29 420	43 994
		...					
		138	0	81	74	6 455	3 165
		139	0	81	57	3 582	4 360
		140	0	81	85	2 932	4 485
		141	0	81	102	129	1 165
...							

Tableau 6.4 Calibrated Coefficients of the Utility Function (Enhanced Destination Ranking Procedure)

<i>Conditional Logit Regression</i>				<i>Pseudo R<sup>2</sup> =35,31%</i>		
<b>Model Parameter</b>	<b>Coefficient</b>	<b>Std. Err.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
<b><math>u_1</math></b>	-0,0289716	0,0001372	-211,19	0,000	-0,0292405	-0,0287027
<b><math>u_2</math></b>	0,000011	3,68e-08	545,85	0,000	0,00002	0,0000202
<b><math>u_3</math></b>	0,0000201	1,99e-07	55,34	0,000	0,0000106	0,0000114

Tableau 6.5 Sample of New Destination Ranks Based on Probabilities (Enhanced Destination Ranking Procedure)

<b>Origin</b>	<b>Destination</b>	<b>PT Trip Duration Between Origin and Each Destination</b>	<b>Probability of Choosing j, When Leaving From i</b>
<i>i</i>	<i>j</i>	<i>d<sub>i_j</sub></i>	<i>p<sub>i_j</sub></i>
...			
102	101	14	48,88%
	102	0	15,05%
	119	53	5,34%
	105	31	3,33%
	107	20	3,14%
	...		
	138	80	0,15%
	135	79	0,14%
	140	90	0,11%
	134	102	0,08%
	141	107	0,06%
...			

## 6.5 Modeling Performance Analysis

In this section, an analysis of the models' performance is presented by means of several global and Goodness-of-Fit measures. Almost all measures confirm that EIOM compared to BIOM behaves better in reproducing the observed OD matrix derived from the 2008 GMA OD survey.

### 6.5.1 Cumulative Number of Trips versus Trip Duration

Figure 6.3 presents the visual similarity of cumulative number of trips versus trip duration for two developed models and the observed values of  $T_{ij}$ . The figure confirms that both models reproduce similar curves to the observed curve. In fact the correlation between EIOM and the actual  $T_{ij}$  curves is 99,6 %.

Based on the figure, BIOM overestimates the cumulative number of trips as trip duration increases. This is probably due to the basic destination ranking procedure that is based only on PT trip duration. By using the enhanced destination ranking procedure, EIOM underestimates the cumulative number of trips as trip duration increases, probably because of more restrictive variables used for ranking the destinations from each origin. We could see that EIOM compared to BIOM yields a closer curve to the observed  $T_{ij}$  curve.

Tableau 6.6 Calibrated Parameters of BIOM and EIOM

Model Parameter	EIOM						BIOM
	Coefficient	Std. Err.	t	P> t	[95% Conf. Interval]		Coefficient
<b>a<sub>1</sub></b>	-0,1641485	0,0340167	-4,83	0,000	-0,2309518	-0,0973452	-0,1985162
<b>a<sub>2</sub></b>	0,176492	0,0095483	18,48	0,000	0,1577408	0,1952432	0,2303826
<b>a<sub>3</sub></b>	-0,0477782	0,0412879	-1,16	0,248	-0,1288609	0,0333045	0,0218092
<b>b</b>	0,0288631	0,0100224	2,88	0,004	0,0091808	0,0485455	0,0203322
<b>P</b>	0,000244%	1,18e-07	20,74	0,000	0,00000221	0,00000267	0,000144%

Table 6.7 presents three measures for an analytical evaluation of similarity between each modeled curve (BIOM or EIOM) versus the observed  $T_{ij}$  curve. Better values for EIOM compared to BIOM confirm that EIOM curve is analytically more similar to the actual  $T_{ij}$  curve.

Tableau 6.7 Similarity Measures for Comparing Modeled Cumulative Number of Trips versus Trip Duration

Similarity Measures	BIOM versus Observed $T_{ij}$	EIOM versus Observed $T_{ij}$
Root of sum of squared errors	240 784	143 285
Mean of error values	19 628	12 637
Standard deviation of error values	12 165	5 378



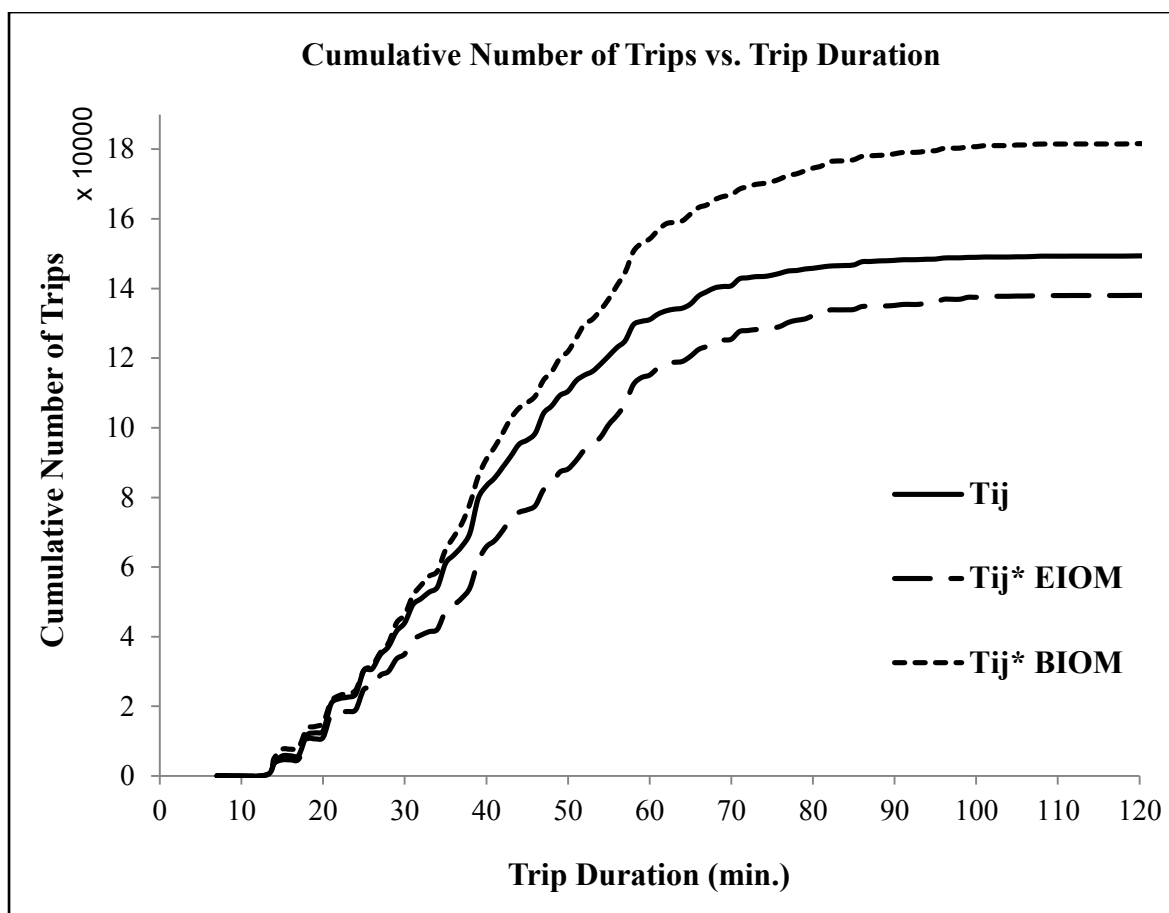


Figure 6.3 Cumulative Number of Trips vs. Trip Duration for BIOM and EIOM

### 6.5.2 Trip Production and Trip Attraction Modeling Performance

Analyzing the performance of trip distribution models in reproducing trip production and attraction values is of great interest. Trip production and attraction measures presented in Table 6.8,  $R^2$  and Dissimilarity Index (DI) confirm that EIOM compared to BIOM yields more similar values to the observed trip production and attraction values.

### 6.5.3 Goodness-of-Fit Measures

Several Goodness-of-Fit measures that compare entries cell-by-cell in the observed and estimated matrices are presented in Table 6.8:

- Mean trip duration error: This is the difference between the mean trip durations estimated by the model and derived from the OD survey. Results confirm that based on this measure, both BIOM and EIOM behave well.
- Coefficient of determination ( $R^2$ ): Although some studies showed that in some cases  $R^2$  may yield artificially high values in Goodness-of-Fit applications, we present it as a traditional measure, because it is one of the most cited measures in the literature (Black, 1991; Knudsen & Fotheringham, 1986; Smith & Hutchinson, 1981; Yaldi et al., 2011). Based on the  $R^2$  values, EIOM is more accurate than BIOM.
- Mean Absolute Error (MAE) and Normalized MAE (NMAE): Larger values of MAE and NMAE represent less accurate model fits (Smith & Hutchinson, 1981). These measures also show that EIOM is more accurate than BIOM.
- Dissimilarity Index (DI) or Percentage Misallocated Error (PME): This measure shows the proportion of flows that are allocated to wrong cells in the matrix (Evans & Pooler, 1987; Hu & Pooler, 2002). As larger values of this measure show larger dissimilarities between the estimated and the observed OD survey matrices, EIOM behaves better than BIOM.
- Root Mean Square Error (RMSE): Smaller value for EIOM confirms that it behaves better than BIOM based on this measure.

Based on all Goodness-of-Fit measures that are presented in Table 6.8, EIOM is more accurate than BIOM. Moreover, EIOM allows policy-makers and analysts to study potential changes in PT trip distribution pattern and also destination attractiveness due to modifications in demography, job locations and also PT supply.

In the next section, we study limitations of BIOM and EIOM from a spatial point of view, to understand their limits.

Tableau 6.8 Goodness-of-Fit Measures

( $T_{ij}$ : Observed number of trips from  $i$  to  $j$ ,  $T_{ij}^*$ : Estimated number of trips from  $i$  to  $j$ ,  
 $T$ : Total number of trips derived from the OD survey,  $N$ : Number of estimated OD pairs)

Performance Analysis Measures			BIOM	EIOM
Global Measures	Mean trip duration (37,5 min observed)		35,7 min	39,4 min
	$T_{ij}=k.T_{ij}^*$ (Value of k)		0,8662	<b>0,8923</b>
	$R^2$		84,88%	<b>88,47%</b>
	MAE	$MAE = \frac{\sum_{i,j}  T_{ij} - T_{ij}^* }{N}$	120,17	<b>97,36</b>
	NMAE	$NMAE = \frac{MAE}{T/N}$	0,37	<b>0,30</b>
	DI (PME)	$DI = \frac{50}{T} \sum_{i,j}  T_{ij} - T_{ij}^* $	24,93	<b>20,32</b>
Trip Production Measures	RMSE	$RMSE = \sqrt{\frac{\sum_{i,j} (T_{ij} - T_{ij}^*)^2}{N}}$	208,08	<b>190,80</b>
	$E_i=k.E_i^*$ (Value of k)		1,1898	<b>0,8805</b>
	$R^2$		<b>95,13%</b>	94,43%
	DI (PME)		12,58	<b>8,98</b>
Trip Attraction Measures	$A_j=k.A_j^*$ (Value of k)		0,9133	<b>0,9706</b>
	$R^2$		91,88%	<b>94,85%</b>
	DI (PME)		19,60	<b>12,41</b>

### 6.5.4 Spatial Limitations of the Models

The spatial limitations of BIOM and EIOM could be presented by analyzing the spatial distribution of underestimation and overestimation residual errors, in other words by means of plotting desire lines between OD pairs that have underestimation or overestimation errors more than 50 %.

Regarding the underestimation residuals with both BIOM and EIOM, few OD pairs yield spatial errors, and also these errors have a random nature.

Concerning the overestimation errors, plotting the desire lines for both models shows less residual errors reported by EIOM compared to BIOM. Moreover, overestimation spatial errors of EIOM are concentrated for OD pairs that destined for a big size zone. This leads us to suppose that probably changing the zoning system and presenting more uniform-sized zones, which is a part of our ongoing research, could resolve this issue.

## 6.6 Conclusion and Future Research

In this study we developed an Enhanced Intervening Opportunities Model (EIOM) for PT trip distribution. The main difference between this model and the Basic Intervening Opportunities Model (BIOM) is in the destination ranking procedure that is used for the model calibration. Both models are calibrated on work trips, with single constraints on trip production for PT morning peak hour trips within the Island of Montreal, Canada. The flow of data that is presented in Figure 6.2 shows different data sources and the way that each required data set is obtained from the raw data. The analysis of modeling performance by means of several global and Goodness-of-Fit measures confirmed that EIOM is well-behaved and more accurate than BIOM.

We studied also limitations of EIOM from a spatial point of view. In other words, by comparing observed and estimated number of trips on an underestimation-overestimation basis, spatial limits of the models are discovered. This could help us to find improvement strategies in terms of explanatory variable choice, zoning system and model formulation modifications, which presents a part of our ongoing and future research.

The presented formulation of EIOM (Equation 6.5) and the enhanced destination ranking procedure represent great potentials for policy-makers and PT analysts due to the sociodemographic and PT supply variables. Also, as all the required data for calibrating EIOM comes from external and independent sources, this model could be used for the sake of PT trip generation and distribution forecasting for the future.

The presence of PT supply characteristics in EIOM allows us to study the effect of potential changes of mobility pattern due to changes in PT supply. This could be of great interest for transport demand “seeding” studies due to new PT projects or urban development.

Also EIOM could be used for studying the mobility pattern of a population due to changes in share of each age cohort. As it was shown in some studies, different sociodemographic groups have different behaviors in a PT network (Nazem et al., 2011), and this is the main reason for considering different sociodemographic groups in the present study. This could be of great interest especially for developed countries with population aging phenomena.

In contrast with the basic procedure, the enhanced destination ranking procedure considers trip purpose in ranking alternative destinations from a given origin. This results different destination ranking lists for different trip purposes, which is more realistic and reasonable.

Regarding trip duration values, the fact that they are used only in the destination ranking procedure, and not explicitly in the model as it was used in the GM, makes BIOM and EIOM less dependent to exact trip duration values. In contrast, it is of a great interest to study the sensitivity of IIOM to trip duration values, because a minor change in trip duration might change the ranking of zones, which will result in changes in number of intervening opportunities.

For further research, we propose a bi-level optimization approach, by using data from an automated fare collection system. At the first level, we will define MS trip production based on sociodemographic, socioeconomic and PTLOS variables, and then we will calibrate it with data derived from smart card validations. The second level is dedicated to the calibration of an EIOM by means of the OD survey data and trip production values calculated at the first level. This new model could be interesting, because it allows us the usage of the detailed data derived from smart card validations. Also it allows the integration of two data sources of OD survey and smart card validation data in the calibration of a single model.

Another research topic is the development of EIOM at a grid level. This could present interesting results for analyzing the sensitivity of PT trip distribution models to the study level. Moreover, we could probably ameliorate spatial errors experienced in the present study at the MS level.

We conclude that EIOM is advantageous due to its behavioral and supply-dependent bases. Also the enhanced destination ranking procedure might be used for behavioral analysis, in terms of choosing alternative destinations based on their PTLOS, number of opportunities and OD trip duration values.

## **Acknowledgments**

The authors acknowledge the support of Natural Sciences and Engineering Research Council of Canada (NSERC) and of the Agence Métropolitaine de Transport (AMT). We thank the survey consortium which provided the Greater Montreal Origin-Destination survey and the GTFS data. Special thanks to François Godefroy, Audrey Godin, Pierre-Léo Mongeon-Bourbonnais, Éric Poliquin, Hubert Verreault, and other members of the Mobilité chair at École Polytechnique de Montréal for their help to obtain and prepare different required data sets.

## **CHAPITRE 7 : RÉSULTATS COMPLÉMENTAIRES**

Ce chapitre vise à présenter quelques résultats complémentaires à ceux discutés aux Chapitres 4, 5 et 6. La première section explique brièvement des raffinements apportés au modèle de choix d'itinéraires en TC présenté au Chapitre 4. Ces raffinements touchent notamment le temps d'accès en TC en distinguant les différents profils sociodémographiques par le biais de la considération des vitesses de marche personnalisées.

La deuxième section consiste à examiner les différentes formalisations du modèle d'opportunités pour la distribution des déplacements tous motifs en TC. Les résultats obtenus nous ont orientés vers la forme du modèle d'opportunités qui a été présentée au Chapitre 5.

### **7.1 Raffinement du modèle de choix d'itinéraires en TC en particulierisant la vitesse de marche pour les différents profils sociodémographiques**

Nous rappelons tout d'abord une hypothèse essentielle pour le modèle de choix d'itinéraires présenté au Chapitre 4 : l'ensemble des itinéraires alternatifs développé à partir du modèle Z (Le Gallou, 2009) est un ensemble non-biaisé et donc les résultats obtenus ne sont pas biaisés en faveur d'un élément spécifique du modèle. Dans le cas où les itinéraires alternatifs sont biaisés, par exemple par la sous-présence ou sur-présence d'un mode particulier, ou en favorisant ou défavorisant un élément particulier dans le modèle, les résultats obtenus seront également biaisés. Afin d'avoir une idée sur les détails du niveau de performance des modèles de choix d'itinéraires, nous présentons brièvement une analyse des itinéraires non-reproduits par le modèle complet. En appliquant les facteurs d'expansion des déplacements selon l'enquête OD 2003, nous obtenons une performance de 53 %, qui est très similaire à la performance de 54 % présentée dans le Tableau 7.6 sans considérer les facteurs d'expansion. Le Tableau 7.1 montre l'écart entre les valeurs observés dans l'enquête OD et celles calculées à partir du modèle complet pour l'ensemble des variables explicatives. Le Tableau 7.1 confirme que les écarts sont acceptables au niveau des variables explicatives, notamment quant aux temps en véhicule et temps d'attente.

Tableau 7.1 Écart entre les valeurs observées et celles modélisées pour les itinéraires non-reproduits par le modèle complet (vitesse de marche constante)

Variable	Écart
NBus	25%
NSub	17%
NBusBus	50%
NSubBus	60%
NBusSub	33%
TBus	10%
TSub	1%
TWaitBus	7%
TWaitSub	14%
TEntWalk	11%
TConWalk	43%
TEgrWalk	13%

Cette section s'inscrit dans l'amélioration du modèle présenté au Chapitre 4. Dans le modèle développé à la section 4.4, plusieurs variables relatives au temps d'accès en TC sont considérées :

- TEntWalk : Temps de marche entre l'origine et le premier point d'embarquement ;
- TConWalk : Temps de marche à la correspondance ;
- TEgrWalk : Temps de marche entre le dernier point de débarquement et la destination.

L'approche utilisée au Chapitre 4 consiste à calculer le temps de marche à partir de la distance de marche et une vitesse constante de 6 km/h, peu importe le profil sociodémographique de la personne qui effectue le déplacement. Cette approche a permis d'atteindre une performance de bon choix d'itinéraire de 54 % et de 62 % pour le modèle complet et le modèle par segments de population, respectivement.

Le raffinement proposé consiste à considérer des vitesses de marche différentes selon les caractéristiques sociodémographiques. Cela personnalise par la suite le temps d'accès pour chaque personne selon sa vitesse de marche. Les vitesses supposées sont présentées au Tableau 7.2 (Bohannon, 1997).

Tableau 7.2 Vitesse de marche selon les profils sociodémographiques

Vitesse de marche (km/h)	Homme	Femme
0-19 ans	5,0148	5,0652
20-64 ans	5,1048	4,9473
65+ ans	4,7880	4,5792

La calibration des modèles complet et par segment à partir des nouveaux temps d'accès a permis de calculer les nouveaux coefficients des fonctions d'utilité. Ces coefficients sont présentés aux Tableaux 7.3, 7.4 et 7.5.

Afin d'apprécier l'effet de ce raffinement sur les coefficients des fonctions d'utilité, nous comparons les Tableaux 4.5 et 7.3 qui représentent les modèles par sexe avec vitesses de marche constante et différentes. Nous pouvons tirer certains constats pour les deux modèles :

- les hommes sont plus sensibles au nombre de lignes de métro empruntées, alors que les femmes sont plus sensibles au nombre de lignes de bus empruntées.
- les femmes sont plus sensibles aux correspondances, notamment celles entre deux modes de TC différents.
- les hommes et les femmes ont des sensibilités similaires quant aux temps dans le véhicule et d'attente pour le bus. Par contre, les femmes sont plus sensibles au temps d'attente pour le métro.
- les hommes sont plus sensibles au temps d'accès, alors que les femmes sont plus sensibles aux temps de marche dans les correspondances.

Les constats confirment que sur une base comparative par sexe, la pénibilité des composants de la fonction d'utilité reste similaire pour les deux modèles avec des vitesses de marche constante et différentes. D'ailleurs, la comparaison des modèles complets aux Tableaux 4.5 et 7.3 montre que les variables de nombre de lignes empruntées (NBus et NSub) et de nombre de correspondances (NBusBus, NSubBus et NBusSub) ont des coefficients avec des valeurs absolues plus élevées en considérant des vitesses de marche différentes. Nous constatons également que toutes les variables de temps (TBus, TSub, TWaitBus, TWaitSub, TEntWalk, TConWalk et TEgrWalk) ont des coefficients avec des valeurs absolues moins élevées en considérant des vitesses de marche différentes.



Globalement l'approche raffinée en distinguant les vitesses de marche permet d'améliorer la performance du modèle de choix d'itinéraires. Cette performance se traduit par une amélioration de pourcentage de prédiction d'itinéraires choisis de 63 % pour le modèle complet à 71 % pour le modèle par segment.

Tableau 7.3 Coefficients de la fonction d'utilité du modèle de choix d'itinéraires en TC par sexe (vitesses de marche différentes)

Variable	Homme	Femme	Modèle complet
<b>NBus</b>	-2,6787	-2,8045	-2,7319
<b>NSub</b>	-0,7154	-0,3741	-0,5358
<b>NBusBus</b>	0,6925	0,9645	0,8361
<b>NSubBus</b>	0,0814	0,4512	0,2823
<b>NBusSub</b>	0,3855	0,8926	0,6629
<b>TBus</b>	0,1973	0,2025	0,1996
<b>TSub</b>	0,2242	0,2283	0,2260
<b>TWaitBus</b>	0,1252	0,1311	0,1286
<b>TWaitSub</b>	-0,1031	-0,2748	-0,1923
<b>TEntWalk</b>	0,0807	0,0665	0,0731
<b>TConWalk</b>	-0,0284	-0,1453	-0,0933
<b>TEgrWalk</b>	0,1156	0,0809	0,0974

Tableau 7.4 Coefficients de la fonction d'utilité du modèle de choix d'itinéraires en TC par cohortes d'âge (vitesses de marche différentes)

Variable	0-19 ans	20-64 ans	65+ ans	Modèle complet
<b>NBus</b>	-1,3698	-3,2419	-1,9007	-2,7319
<b>NSub</b>	-0,5396	-0,4405	-2,0778	-0,5358
<b>NBusBus</b>	-0,4805	1,2657	0,1587	0,8361
<b>NSubBus</b>	-0,6433	0,6072	0,2285	0,2823
<b>NBusSub</b>	-0,1439	0,9797	0,0808	0,6629
<b>TBus</b>	0,1985	0,2015	0,2694	0,1996
<b>TSub</b>	0,2288	0,2226	0,2613	0,2260
<b>TWaitBus</b>	0,1140	0,1350	0,2168	0,1286
<b>TWaitSub</b>	-0,0364	-0,2649	0,4733	-0,1923
<b>TEntWalk</b>	0,0898	0,0634	0,1331	0,0731
<b>TConWalk</b>	0,0180	-0,1230	-0,1185	-0,0933
<b>TEgrWalk</b>	0,1231	0,0865	0,0805	0,0974

La comparaison de performances des modèles développés présentée au Tableau 7.6 confirme l'avantage de considérer les segments de population ainsi que de personnaliser les vitesses de marche en lien avec les segments de population concernés.

Tableau 7.5 Coefficients de la fonction d'utilité du modèle de choix d'itinéraires en TC par cohortes d'âge et par sexe (vitesses de marche différentes)

Variable	0-19 ans		20-64 ans		65+ ans		Modèle complet
	Homme	Femme	Homme	Femme	Homme	Femme	
<b>NBus</b>	-0,9237	-1,8018	-3,4548	-3,1215	-2,1529	-2,2015	-2,7319
<b>NSub</b>	-0,8873	-0,2031	-0,5757	-0,3488	-4,0960	-0,0372	-0,5358
<b>NBusBus</b>	-0,8502	-0,1317	1,2389	1,2824	0,4616	0,5638	0,8361
<b>NSubBus</b>	-0,8205	-0,4596	0,4730	0,7053	-0,2742	0,9709	0,2823
<b>NBusSub</b>	-0,6888	0,3691	0,8698	1,0701	-0,4376	1,1475	0,6629
<b>TBus</b>	0,1955	0,2037	0,2022	0,2023	0,2681	0,2663	0,1996
<b>TSub</b>	0,2355	0,2255	0,2173	0,2269	0,2409	0,2772	0,2260
<b>TWaitBus</b>	0,1060	0,1217	0,1351	0,1351	0,2851	0,1653	0,1286
<b>TWaitSub</b>	0,1278	-0,2026	-0,2055	-0,3087	1,4875	-0,6435	-0,1923
<b>TEntWalk</b>	0,0907	0,0896	0,0722	0,0559	0,1503	0,0991	0,0731
<b>TConWalk</b>	0,0341	-0,0047	-0,0351	-0,1866	-0,3420	-0,0806	-0,0933
<b>TEgrWalk</b>	0,1336	0,1127	0,1077	0,0694	0,1035	0,0617	0,0974

Tableau 7.6 Comparaison de performances des modèles de choix d'itinéraires en TC

<b>Performance des modèles au niveau de pourcentage de prédiction d'itinéraire choisi observé</b>	Vitesse de marche constante pour l'ensemble de la population (6 km/h)	Vitesses de marche différentes pour les segments sociodémographiques (voir le Tableau 7.2)
Modèle complet	54%	63%
Modèle par segments sociodémographiques	62%	71%

## 7.2 Formalisation d'un modèle d'opportunités pour la distribution des déplacements tous motifs en TC

Différentes formalisations du modèle d'opportunités pour la distribution des déplacements tous motifs sont examinées afin de choisir la formalisation la plus performante. Nazem et al. ont présenté le développement des MFO, MHO et un modèle gravitaire (Nazem et al., 2012), et dans cette section, le MIO ainsi qu'une comparaison des performances de tous ces modèles sont brièvement présentés.

Le MIO est nommé intégré parce qu'on calibre un modèle non linéaire en intégrant tous les motifs de déplacements dans un seul niveau. Le nombre de déplacements entre les zones  $i$  et  $j$  pour chaque motif est présenté comme suit :

$$T_{ij} = E_i \cdot \frac{e^{-P \cdot O_{j-1}} - e^{-P \cdot O_j}}{1 - e^{-P \cdot O_j}}, \forall i, j \quad (7.1)$$

$T_{ij}$  : Nombre de déplacements entre les zones  $i$  et  $j$

$E_i$  : Nombre total de déplacements produits par la zone  $i$

$P$  : Probabilité de choisir une opportunité dans une zone

$O_j$  : Nombre d'opportunités cumulé entre les zones  $i$  et  $j$ , incluant  $j$

$O_{j-1}$  : Nombre d'opportunités cumulé entre les zones  $i$  et  $j$ , excluant  $j$

$O_j$  : Nombre total d'opportunités

Et le nombre de déplacements tous motifs entre  $i$  et  $j$  est :

$$T_{ij}^{Total} = k_T \cdot T_{ij}^{Travail} + k_E \cdot T_{ij}^{Étude} + k_A \cdot T_{ij}^{Autre}, \forall i, j \quad (7.2)$$

Après avoir préparé les données nécessaires, le MIO est calibré dans le logiciel d'analyses statistiques STATA, à l'aide d'un processus non linéaire. Les résultats présentés au Tableau 7.7 montrent que tous les paramètres du modèle sont statistiquement significatifs avec un intervalle de confiance de 95 %.

Le Tableau 7.8 confirme que le MIO par rapport au MFO, MHO et modèle gravitaire représente mieux la réalité, d'où l'idée de développement du modèle intégré de génération-distribution présenté au Chapitre 5.

Tableau 7.7 Paramètres de MIO (déplacements tous motifs)

Motif de déplacement	$P_{MIO}$	$k_{MIO}$
Travail	0,000152%	1,679834
Étude	0,033890%	0,213729
Autre	0,000051%	1,247547

Tableau 7.8 Mesures statistiques de validité de l'ajustement des MFO, MHO, MIO et modèle gravitaire (déplacements tous motifs)

		Modèles d'opportunités			Modèle gravitaire
		MFO	MHO	MIO	
Mesures globales	Erreur de temps de parcours moyen	1,63%	<b>0,04%</b>	0,90%	4,15%
	$T_{ij}^* = k \cdot T_{ij}$ (Valeur de k)	0,874	0,886	<b>0,896</b>	0,836
	$R^2$	85,0%	88,1%	<b>89,2%</b>	80,95%
	PHI	98 004	94 253	<b>87 011</b>	94 504
	NPHI	0,408	0,393	<b>0,373</b>	0,394
	MAE	136,35	128,03	<b>112,16</b>	153,32
	NMAE	0,422	0,396	<b>0,347</b>	0,474
	DI	21,11	19,82	<b>17,88</b>	23,73
	RMSE	249,6	214,4	<b>199,4</b>	250,2
Mesures de génération de déplacements	$E_i^* = k \cdot E_i$ (Valeur de k)	1,083	1,074	<b>1,031</b>	1,110
	$R^2$	99,3%	98,6%	97,8%	<b>99,7%</b>
	DI	<b>5,24</b>	5,84	5,94	5,54
Mesures d'attraction de déplacements	$A_j^* = k \cdot A_j$ (Valeur de k)	0,865	<b>0,918</b>	0,906	0,912
	$R^2$	91,4%	94,8%	<b>95,6%</b>	92,8%
	DI	13,52	11,77	<b>11,77</b>	12,81

Le chapitre suivant est consacré à une discussion générale sur l'ensemble des résultats obtenus au cours de la recherche en lien avec la problématique et les objectifs ciblés au début de la thèse.

## **CHAPITRE 8 : DISCUSSION GÉNÉRALE**

L'objectif de cette thèse est de contribuer à la modélisation des déplacements en TC en bonifiant des modèles déjà développés ou en présentant de nouvelles façons d'examiner la problématique.

La revue de littérature relative aux différentes approches de modélisation et les limites associées à chacune ont permis d'identifier des pistes d'intervention pour bonifier le processus de la modélisation en développant de nouvelles formalisations ou en profitant de nouvelles données en TC. Nous rappelons les questions de départ en lien avec les limites identifiées lors de la revue de littérature :

- Comment mieux considérer les caractéristiques sociodémographiques et socioéconomiques de la population dans les modèles de prévisions de déplacements notamment dans les modèles de distribution et de choix d'itinéraires en TC ?
- Comment estimer la demande de TC en fonction des changements apportés à l'offre à un niveau agrégé surtout pour des fins de prévisions de mobilité à long terme ?
- Comment analyser le choix de destination des usagers de TC à l'aide d'une approche d'analyse comportementale, spécifiquement un modèle d'opportunités ?

Cette recherche a permis de contribuer à la modélisation des déplacements en TC sous plusieurs aspects :

- Bien que les caractéristiques socioéconomiques soient prises en compte dans les modèles de génération de déplacement, elles ne sont néanmoins pas généralement considérées dans les modèles de distribution et de choix d'itinéraires en TC. Cette recherche permet une réflexion plus réelle à cet égard en développant un modèle de choix d'itinéraires par segments sociodémographiques des usagers du TC. Cette meilleure compréhension au niveau du comportement de mobilité pourrait être considérée dans plusieurs processus de modélisation en TC dont l'approche totalement désagrégée de MADITUC utilisée dans la région montréalaise. Par ailleurs, les constats relatifs aux différences comportementales tirés de ce modèle nous ont orientés vers un choix de variables plus fin au niveau de la population pour le développement de modèles d'opportunités.

- Les modèles d'opportunités développés permettent une meilleure compréhension des comportements de mobilité, grâce à leurs caractères agrégés.
- Le modèle intégré d'opportunités pour la génération-distribution des déplacements en TC a du potentiel d'application aux études de prévisions de la demande en TC à long terme, en raison de l'ajout de variables sociodémographiques (population par trois cohortes d'âge), socioéconomiques (revenu moyen des personnes), et de l'offre de TC (niveau de service de TC). Ces variables sont prévisibles ou supposables dans le temps, d'où l'avantage de ce modèle pour évaluer divers scénarios de développement urbain ainsi que l'impact de nouveaux projets de TC sur le comportement de mobilité.
- La formalisation de modèles d'opportunités a permis d'utiliser les données provenant de multiples sources dont l'enquête OD, le recensement et les fichiers GTFS. Les diagrammes de flux de données présentés pour chaque modèle contribuent à identifier le potentiel d'insérer de nouvelles données dans le processus de la modélisation, tout en essayant de minimiser les interdépendances entre les données entrantes (Figures 5.2, 6.2 et B.5).

Certains aspects nécessitent plus d'approfondissement méthodologique et technique. En voici quelques-uns :

- Choix du territoire d'étude

La réalisation des études en transport nécessite généralement la détermination d'un territoire et un ensemble de population pour étudier leur comportement de mobilité. Il semble que le niveau de captivité ait une influence sur la validité de certains modèles en TC, d'où l'importance de bien choisir le territoire d'étude (Varley & Chen, 2010).

Les développements méthodologiques au cours de cette recherche sont examinés sur le territoire de l'île de Montréal (Québec), grâce à la disponibilité des données requises et pour la complexité technique, sachant que la méthodologie est généralisable à d'autres territoires.

- Choix du niveau de résolution d'étude

Cela s'applique plus particulièrement à l'approche agrégée de la modélisation des déplacements. Le choix de niveau de résolution est fortement dépendant de l'objectif d'étude et du niveau de précision nécessaire.

Dans cette étude, nous avons utilisé les Secteurs Municipaux (SM) de l'île de Montréal (Québec) comme unité d'analyse pour les développements méthodologiques relatifs au modèle d'opportunités. Cette approche est généralisable à d'autres niveaux de résolution plus fins en faisant attention à la significativité des données disponibles. Dans le chapitre suivant, nous proposons brièvement un niveau de résolution de cellules personnalisables pour des recherches futures.

- Choix de variables représentant le niveau de service du TC

Dans cette étude, nous avons utilisé deux variables agrégées pour représenter le niveau de service du TC : le nombre total de passage-arrêts des véhicules de TC durant une journée et la densité spatiale des arrêts dans chaque zone. Ces deux variables sont calculables à partir des données de GTFS, peu importe le niveau de résolution choisi. L'intérêt d'utilisation de ces variables est qu'elles représentent à la fois les caractéristiques opérationnelles de service de TC fourni et les infrastructures telles que la distribution spatiale des arrêts.

L'utilisation de ces variables est bénéfique dans un modèle de prévision de la mobilité à long terme puisque qu'on peut en faire une prévision plausible selon les projets futurs en TC. Plusieurs raffinements peuvent être amenés à ce sujet dont la pondération des variables calculées en fonction des modes de TC, etc.

- Arrimage des méthodologies de modélisation et des nouvelles données

La revue de littérature a démontré des besoins de données différents pour le développement de chaque modèle en transport. L'existence de nouvelles données opérationnelles et passives ouvre le potentiel de les arrimer avec les processus actuels de la modélisation, qui nécessite plus de réflexions méthodologiques et techniques.

Par exemple, les données provenant des systèmes de paiement par CAP sont des données fines et précises mais parfois difficilement exploitables. Cette complexité est d'une part au niveau méthodologique attribuable à l'absence de certaines informations essentielles dans les enregistrements de CAP et d'autre part au niveau technique en raison de la quantité de données importante. Le modèle d'opportunités développé dans cette recherche a un certain potentiel pour exploiter ces données, attribuable à sa formalisation mathématique. Plus précisément pour la calibration de ce modèle contrairement à un MG, le nombre d'attraction des déplacements dans chaque zone n'est pas explicitement requis. Cela représente une piste exploratoire pour

l'utilisation des données de CAP dans la modélisation des déplacements en TC vu que dans la plupart des cas, les points de débarquement ne sont pas connus à cause des systèmes de simple-validation uniquement aux points d'embarquement.

Un autre sujet d'intérêt est le potentiel des données d'horaires planifiés provenant des GTFS pour caractériser l'offre de TC. Dans cette recherche, nous avons utilisé ces données pour représenter le niveau de service dans une zone. Une autre piste est l'utilisation de ces données dans le processus de la codification d'un réseau complexe de TC. L'intérêt général est que ces données suivent la variabilité des services de TC fournis avec des niveaux de précision temporels et spatiaux élevés. Cela représente également un grand avantage pour migrer d'un environnement de simulation en régime permanent vers un environnement en régime dynamique.

Plusieurs pistes d'arrimage de données nouvellement disponibles et des processus de modélisation en TC sont envisageables et s'inscrivent dans une approche de modélisation qu'on pourrait appeler « *approche orientée vers les données* ». Les diagrammes de flux de données peuvent agir comme outil pour mieux apprécier la nature des données nécessaires ainsi que l'interdépendance entre elles (Hato & Kitamura, 2008; Liao & X. Liu, 2010; Nazem et al., 2012, 2013).

Plusieurs autres questions se posent à cet égard : comment définir un cadre global pour l'utilisation des données de types différents dans le processus de la planification en TC ? Quels en sont les potentiels et les limites ? Ces questions soulèvent plus globalement les besoins de développement de nouvelles approches de fusion de données en TC.

- Opérationnalisation des modèles développés

Les chercheurs continuent à développer de nouveaux modèles dans certains cas très complexes. La question qui se pose est la suivante : dans quelle mesure peut-on bénéficier de ces nouvelles avancées dans la pratique ? Plus globalement cela s'inscrit dans l'opérationnalisation des modèles en TC, un sujet d'intérêt majeur pour les professionnels. La complexité relative à l'application des modèles aux cas réels est généralement associée à deux aspects différents : la disponibilité des données requises et les complexités méthodologiques et techniques d'application.

Dans le cas des développements présentés dans cette thèse, quelques pistes d'opérationnalisation sont envisageables. En ce qui concerne le modèle de choix d'itinéraires en TC selon les caractéristiques sociodémographiques présenté aux Chapitres 4 et 7, les constats sont intéressants à intégrer dans les modèles de simulation des réseaux de TC, ou bien dans les calculateurs de



trajet commerciaux de type Google Maps. Quant au modèle intégré d'opportunités pour la génération-distribution des déplacements présenté aux Chapitres 5, 6, et 7, plusieurs applications sont envisageables dont son utilisation lors des études d'ensemencement de la demande dans le cadre des projets majeurs de TC.

Les développements méthodologiques de cette thèse peuvent être utiles par différentes AOT à travers du Canada, car les données requises sont des données généralement disponibles surtout dans le cas des grandes villes canadiennes.

- Transférabilité des modèles développés dans le temps et l'espace

La question de la transférabilité des modèles de transport dans le temps et l'espace pour évaluer l'applicabilité des modèles développés pour un cas précis à d'autres cas d'étude est un sujet d'intérêt majeur. Dans le cas de cette recherche, les développements méthodologiques sont testés sur le territoire de l'île de Montréal (Québec), et pour les déplacements effectués durant les heures de pointe du matin. Certaines questions se posent en lien avec ceci : peut-on utiliser la même méthodologie sur d'autres territoires, par exemple sur la grande région de Montréal, et sur d'autres périodes temporelles, par exemple celle hors-pointe ? Quelles sont les modifications nécessaires à apporter dans les modèles ? Quelle est la performance des modèles dans ces cas-là ?

Les résultats obtenus confirment les contributions apportées à certains aspects de la modélisation en développant des modèles plus réalistes et plus opérationnels en lien avec les enjeux actuels et émergents de la mobilité ainsi que des nouvelles données disponibles en TC. Le dernier chapitre présente la conclusion ainsi que certaines limitations et perspectives pour les recherches futures.

## **CHAPITRE 9 : CONCLUSION**

Cette recherche a comme objectif de contribuer à la modélisation des déplacements en TC par le biais de la bonification de la construction théorique et de la formalisation mathématique de multiples modèles en TC, afin qu'ils représentent mieux les phénomènes réels de la mobilité des personnes. La conclusion englobe une synthèse des travaux réalisés, les contributions, les limitations et certaines perspectives.

### **9.1 Synthèse des travaux**

Les travaux présentés dans cette thèse s'articulent autour de plusieurs pistes d'amélioration de la modélisation en TC, en portant une attention particulière aux aspects comportementaux et aussi en utilisant des données nouvellement disponibles. Les travaux touchent plusieurs aspects :

- la considération des caractéristiques sociodémographiques dans la procédure de choix d'itinéraires en TC : dans le Chapitre 4, nous avons segmenté la population en six groupes basés sur les cohortes d'âge et le sexe. Par la suite, pour chaque paire OD un ensemble de choix d'itinéraires est créé et des modèles de choix discret sont calibrés sur chaque segment sociodémographique pour modéliser l'itinéraire choisi observé. Également, un modèle de choix d'itinéraires nommé le modèle complet est calibré sur la population comme un ensemble monolithique. La comparaison de performances des modèles par segment et celui complet confirme que la segmentation sociodémographique de la population permet d'avoir une meilleure représentation de la réalité quant au choix d'itinéraires en TC. Certaines bonifications notamment au niveau de raffinement du calcul du temps d'accès en supposant des vitesses de marche différentes selon les caractéristiques sociodémographiques des passagers sont également présentées au Chapitre 7. Les résultats confirment que la considération des vitesses de marches différentes améliore la performance des modèles de choix d'itinéraires en TC.
- le développement d'un modèle intégré d'opportunités pour la génération-distribution des déplacements en TC : quant au développement de modèle d'opportunités pour tous motifs, trois formalisations différentes, nommées modèles fondamental, hiérarchique et intégré sont examinées dans le Chapitre 7. La comparaison de performances des trois

modèles d'opportunités ainsi qu'un modèle gravitaire confirme l'avantage du modèle intégré. À partir de ces résultats, un modèle intégré d'opportunités pour la génération-distribution des déplacements en TC en ajoutant les variables sociodémographiques (population par trois cohortes d'âge), socioéconomiques (revenu moyen des personnes) et l'offre de TC (niveau de service dérivé des données GTFS) est développé et présenté au Chapitre 5. Ce modèle peut être utilisé en tant qu'un outil de modélisation prospective grâce à la nature prévisible ou supposable de ses variables indépendantes.

- le développement d'un sous-modèle de choix de destination dans l'optique d'amélioration des modèles d'opportunités : ce sous-modèle de choix de destination dont bénéficie le développement des modèles d'opportunités est présenté au Chapitre 6. Principalement ce sous-modèle est utilisé dans une procédure améliorée d'ordonnancement des destinations pour le développement du modèle d'opportunités. L'avantage de cette procédure est entre autres la diminution de la sensibilité du modèle d'opportunités aux changements mineurs de temps de parcours en TC, ainsi que l'amélioration de la performance du modèle d'opportunités.

Au cours de la recherche, d'autres aspects de la modélisation sont également revus notamment en lien avec de nouvelles données en TC. Les résultats obtenus permettent d'avoir une réflexion plus profonde face à certaines problématiques et de considérer de nouvelles opportunités dans la pratique de la modélisation. Voici quelques sujets d'intérêt majeur :

- importance d'intégrer la composition démographique dans le processus de modélisation : le vieillissement et plus généralement le changement dans la composition démographique de la population surtout dans les pays développés et la nécessité de prendre ces changements en considération dans les études de mobilité notamment à long terme ;
- nécessité de procéder à des analyses prospectives de la mobilité à un niveau agrégé dans le cadre des projets majeurs en TC : l'étude de comportements de mobilité à un niveau agrégé, plus spécifiquement la génération et la distribution des déplacements attribuables aux projets en TC afin d'avoir une aperçu agrégé quant à la mobilité avant de procéder à des analyses détaillées ;
- potentiels des nouvelles données en TC : la disponibilité des nouvelles données, entre autre celles planifiées (dérivées des fichiers GTFS) et passives (provenant des systèmes

de paiement par CAP et des systèmes de localisation des véhicules de TC), et l'opportunité de les intégrer dans le processus de la modélisation ;

- importance de l'opérationnalisation des modèles en TC : les enjeux relatifs à la mise en pratique des modèles en TC en lien avec le manque de données requises ainsi que les complexités méthodologiques et techniques d'application de certains modèles développés par les chercheurs dans la pratique courante des professionnels de transport.

## 9.2 Contributions

Les résultats obtenus au cours de cette recherche contribuent à la modélisation des déplacements en TC sous plusieurs aspects et applications. En voici quelques-uns :

- le modèle de choix d'itinéraires en fonction des caractéristiques sociodémographiques peut être intégré dans les fonctions d'impédance utilisées par les logiciels de modélisation tels que MADIGAS.
- le modèle intégré d'opportunités permet de prévoir la génération et la distribution des déplacements en TC en fonction des caractéristiques sociodémographiques (population par trois cohortes d'âge), socioéconomiques (revenu moyen des personnes), et l'offre de TC (niveau de service). L'avantage de ce modèle est qu'il peut considérer des changements apportés à ces éléments pour prévoir les matrices OD des déplacements pour le futur.
- le sous-modèle de choix de destination de nature choix discret permet entre autre de mieux comprendre la structure de choix de destination des usagers du TC.

## 9.3 Limitations

Bien que les résultats obtenus dans cette thèse aient permis de contribuer à l'avancement de la modélisation en TC, quelques limitations sont présentes. Dans les Chapitres 4, 5, et 6 nous avons mentionné certaines limitations spécifiques aux modèles développés. À partir de l'ensemble des résultats issus des articles et d'autres analyses complémentaires, quelques limitations générales sont constatées :

- certaines limitations spatiales résultent de choix de niveau de résolution au départ, notamment dans le cas des études agrégées. La définition d'un niveau de résolution approprié à chaque étude en lien avec ses objectifs et son niveau de précision nécessaire est essentielle. La reconsidération de cette problématique par le biais d'un système spécifique de découpage zonal est discutée dans les perspectives.
- la plupart des modèles sont développés pour la période de pointe du matin, du fait de l'importance particulière de cette période temporelle pour la mobilité en TC. Par ailleurs, les outils de modélisation comprenant les réseaux de simulation et les paramètres associés sont généralement spécifiques à cette période. Cela représente une limite pour les planificateurs de TC surtout en lien avec les enjeux relatifs à la mobilité hors-pointe de certains segments de population utilisateurs de TC, notamment les personnes âgées dont la proportion est en pleine croissance. Nous présentons quelques pistes de réflexion face à cette limitation temporelle dans les perspectives.
- le manque de données nécessaires et dans certains cas la complexité technique de traitement des données représentent des limitations majeures qui méritent une attention particulière pour les développements futurs. Par exemple dans le cas de développement de modèle de choix d'itinéraires en considérant les segments sociodémographiques, le nombre d'observation provenant de l'enquête OD pour certains segments de la population notamment les personnes âgées est faible. Cela peut affecter la significativité des résultats ou de certains coefficients des modèles développés. Une piste exploratoire face à cette limitation peut être la valorisation des données passives de type CAP pour la caractérisation de la demande en transport.

## 9.4 Perspectives

Les résultats obtenus et les limitations générales et spécifiques constatées dans cette thèse permettent de proposer quelques perspectives pour les recherches futures.

Des perspectives spécifiques pour chaque modèle présenté dans les Chapitres 4, 5, et 6 sont discutées dans ces chapitres respectivement. Voici quelques perspectives générales de cette recherche :

- Reconsidération de niveau de résolution dans le processus de la modélisation

Le niveau de résolution choisi pour le développement d'un modèle agrégé en TC a des effets sur la performance du modèle ainsi que sur la précision des résultats obtenus. Pour les recherches futures dans le domaine des modèles agrégés, nous proposons une modélisation au niveau d'une grille uniforme composée de cellules identiques. L'intérêt de cette approche d'analyse par cellule est que l'analyste pourra regrouper les cellules selon certains critères donnés pour des objectifs spécifiques. Également l'analyste est en mesure d'examiner la sensibilité des résultats à la taille des cellules, dépendamment de la précision des données disponibles.

- Reconsidération de formalisations mathématiques de certains modèles

Plusieurs nouvelles formes de modèles sont présentées dans les Chapitres 4, 5, 6, et 7. De façon générale, nous proposons de reconsidérer certaines formalisations actuelles des modèles pour en tirer les potentiels d'arrimage avec les nouvelles données en TC. De plus, les nouvelles formalisations des modèles doivent permettre de considérer des nouveaux enjeux et aspects comportementaux actuels. Par exemple dans le cas des modèles de choix d'itinéraires, un composant intéressant dans la fonction d'utilité peut être la notion de « *confort* ».

- Valorisation des nouvelles données en lien avec le processus de la modélisation

L'objectif est d'intégrer les nouvelles données dans les analyses descriptives et aussi dans la modélisation en TC. Cela pourra être une piste intéressante pour répondre à quelques limitations constatées au cours de cette recherche, notamment quant au manque de nombre d'observations suffisant pour des fins de calibration pour certains segments de population ou certaines périodes temporelles. En termes d'exemple, afin de modéliser les déplacements durant une journée entière ou bien d'autres périodes spécifiques de la journée à part la période de pointe du matin traditionnellement considérée, nous pouvons modéliser l'offre à l'aide des données d'horaires planifiés de GTFS et la demande de TC à l'aide des données passives provenant des systèmes de paiement par CAP. Bien que plusieurs subtilités méthodologiques et techniques soient associées à cette approche, elle présente une opportunité pour la modélisation des déplacements en TC.

L'objectif ultime de la modélisation des déplacements en TC est de représenter les phénomènes réels relatifs à la mobilité des personnes par le biais des modèles mathématiques. Cela nécessite d'une part une meilleure compréhension de comportements de mobilité de la population utilisatrice de TC, et d'autre part le développement des modèles capables de considérer les enjeux

émergents dans le domaine de transport. Les nouvelles données opérationnelles et passives en TC sont un atout pour répondre à certains besoins. Étant donné la nature appliquée de la science de transport, l'opérationnalisation des développements méthodologiques reconnus et validés par la communauté scientifique est un élément d'intérêt majeur pour les professionnels et planificateurs des transports.

## BIBLIOGRAPHIE

- Afandizadeh, S., & Hamedani, S. M. Y. (2012). A fuzzy intervening opportunity model to predict home-based shopping trips. *Canadian Journal of Civil Engineering*, 39(2), 203-222. doi: 10.1139/111-097
- Akwawua, S., & Pooler, J. (2001). The Development of an Intervening Opportunities Model with Spatial Dominance Effects. *Journal of Geographical Systems*, 3, 69-86.
- Almeida, L. M. W., & Goncalves, M. B. (2001). A Methodology to Incorporate Behavioral Aspects in Trip Distribution Models with an Application to Estimate Student Flow. *Environment and Planning A*, 33, 1125-1138.
- AMT. (2010). Mobilité des Personnes. Montreal, QC: Secrétariat à l'Enquête Origine-Destination (OD).
- Arifn, Z. N., & Axhausen, K. W. (2012). Investigating Commute Mode and Route Choice Variabilities in Jakarta Using Multi-Day GPS Data. *International Journal of Technology*, 3(1), 45-55.
- Atkins, S. T. (1976). Transportation Planning Models - What the Papers Say. *Traffic Engineering and Control*, 27(9), 460-467.
- Babazadeh, A., Khodakarami, M. A., & Zokaei Ashtiani, H. (2010). *Passenger Origin-Destination Matrix Estimation Employing the Path-Based Formulation of Transit Assignment Problem*. Paper presented at the 89th Annual Meeting of the Transportation Research Board.
- Bachelet, B., & Yon, L. (2005). *Enhancing Theoretical Optimization Solutions by Coupling with Simulation*. Paper presented at the 1st Open International Conference on Modeling and Simulation, Université Blaise Pascal, Clermont-Ferrand, France.
- Bagchi, M., & White, P. (2004). What role for smart-card data from bus systems? *Municipal Engineer*, 157(ME1), 39-46.
- Beimborn, E. A. (2006). A Transportation Modeling Primer. Wisconsin, Milwaukee: Center for Urban Transportation Studies.
- Berkow, M., El-Geneidy, A. M., Bertini, R. L., & Crout, D. (2009). *Beyond Generating Transit Performance Measures: Visualizations and Statistical Analysis Using Historical Data*. Paper presented at the 88th Annual Meeting of the Transportation Research Board, Washington DC.



- Bertini, R. L., & El-Geneidy, A. M. (2003). Generating Transit Performance Measures with Archived Data. *Transportation Research Record*, 109-119.
- Black, W. R. (1991). A Note on the Use of Correlation Coefficients for Assessing Goodness-of-fit in Spatial Interaction Models. *Transportation*, 18(3), 199-206.
- Blogg, M., Semler, C., Hingorani, M., & Troutbeck, R. (2010). *Travel Time and Origin-Destination Data Collection Using Bluetooth Mac Address Readers*. Paper presented at the Australasian Transport Research Forum, Canberra, Australia.
- Blum, J. J., Sridhar, A., & Mathew, T. V. (2010). *Origin-Destination Matrix Generation from Boarding-Alighting and Household Survey Data for a Large Multi-modal Public Transit System*. Paper presented at the 89th Annual Meeting of the Transportation Research Board.
- Bohannon, R. W. (1997). Comfortable and Maximum Walking Speed of Adults Aged 20-79 Years: Reference Values and Determinants. *Age and Ageing*, 26, 15-19.
- Bonnell, P. (2004). *Prévoir la Demande de Transport*. Paris, France: Presses Ponts et Chaussées.
- Bonnell, P., Hombourger, E., Olteanu-Raimond, A.-M., & Smoreda, Z. (2014). *Analyse du potentiel des données « passives » de la téléphonie mobile pour produire des matrices origines-destinations*. Paper presented at the Colloque annuel de la chaire Mobilité, Montréal, Québec.
- Bovy, P. H. L., & Hoogendoorn-Lanser, S. (2005). Modelling Route Choice Behaviour in Multi-modal Transport Networks. *Transportation Journal*, 32, 341-368.
- Boyce, D., O'Neill, C. R., & Scherr, W. (2008). *New Results on Solving the Sequential Travel Forecasting Procedure with Feedback*. Paper presented at the TRB 2008 Annual Meeting, Washington, D.C.
- Caceres, N., Wideberg, J., & Benitez, F. (2007). Deriving Origin Destination Data from a Mobile Phone Network. *Intelligent Transport Systems*, 1(1), 15-26.
- Caussidier, C. (2007). La démarche scientifique dans les différentes disciplines. [http://www.irem.univ-montp2.fr/IMG/pdf/La\\_demarche\\_scientifique-3.pdf](http://www.irem.univ-montp2.fr/IMG/pdf/La_demarche_scientifique-3.pdf)
- Ceder, A. A. (2007). *Public Transit Planning and Operation: Theory, modeling and practice*: Elsevier.

- Celik, H. M. (2010). Sample size needed for calibrating trip distribution and behavior of the gravity model. *Journal of Transport Geography*, 18(1), 183-190. doi: <http://dx.doi.org/10.1016/j.jtrangeo.2009.05.013>
- Cevallos, F., Wang, X., Gan, A., & Chen, Z. (2010). *Improving Transit On-Time Performance with AVL Data: A Timetable Approach*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- Chapleau, R. (1986). *Transit Network Analysis and Evaluation with a Totally Disaggregate Approach*. Paper presented at the World Conference on Transportation Research, Vancouver, Canada.
- Chapleau, R. (2002). *Multiple Uses of Large-scale Origin-Destination Surveys, with a Totally Disaggregate Approach: The Montreal Experience*. Paper presented at the 4th Transportation Specialty Conference of the Canadian Society for Civil Engineering, Montréal, Canada.
- Chapleau, R., & Allard, B. (2010). *Merging AFC, APC, GPS and GIS-T Data to Generate Productivity Indicators and Travel Demand Models in Public Transit*. Paper presented at the 12th WCTR, Lisbon, Portugal.
- Chapleau, R., Allard, B., Trépanier, M., & Morency, C. (2001). Les Logiciels d'Enquête Transport comme Instruments Incontournables de Planification Analytique. *Recherche, Transport, Sécurité*, 70, 59-77.
- Chapleau, R., & Lavigne, P. (1996). *Transport et Emploi dans la Grande Région de Montréal selon la Méthode Orientée-Objet de MADITUC et les Enquêtes OD*. Montréal, Québec: Groupe MADITUC, École Polytechnique de Montréal.
- Chapleau, R., & Morency, C. (2005). *Dynamic Spatial Analysis of Urban Travel Survey Data using GIS*. Paper presented at the 25th Annual ESRI International User Conference, San Diego, California.
- Chapleau, R., & Trépanier, M. (1997). *Transportation Object-Oriented Modeling: an Extension of the Totally Disaggregate Approach*. Paper presented at the International Association of Travel Behavior Research 8th Meeting, Austin, Texas.
- Chapleau, R., Trépanier, M., & Allard, B. (1998). *Practical Implementations of Object-oriented GIS-T*. Paper presented at the World Conference on Transportation Research, Antwerpen, Belgium.

- Chaudhari, J., & Ye, Z. (2010). *Use of Geographical Information System as a Sketch Plan Tool to Replace Traditional Transit Route Planning Practice for College and University Communities*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- Chen, C., Gong, H., Lawson, C., Bialostozky, E., & Muckell, J. (2010). *Evaluating the Feasibility of a Passive Travel Survey Collection in a Complex Urban Environment: A Case Study in New York City*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- Chen, X., Yu, L., Zhang, Y., & Guo, J. (2010). *Analyzing Urban Bus Service Reliability at the Stop, Route and Network Level*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- Cheng, G., Wilmot, C., & Baker, E. (2011). Dynamic Gravity Model for Hurricane Evacuation Planning. *Transportation Research Record: Journal of the Transportation Research Board*, 2234(-1), 125-134. doi: 10.3141/2234-14
- Chow, L.-F., Zhao, F., Li, M.-T., & Li, S.-C. (2005). Development and Evaluation of Aggregate Destination Choice Models for Trip Distribution in Florida. *Transportation Research Record*(1931), 18-27.
- Chu, K. K. A. (2014). *Characterizing Long-Term Transit Service Consumption and Fare Usage Patterns with Disaggregate Smart Card Data*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Chu, K. K. A., & Chapleau, R. (2010). *Augmenting Transit Trip Characterization and Travel Behavior Comprehension with Multi-day Location-Stamped Smart Card Transactions*. Paper presented at the 89th Annual Meeting of the Transportation Research Board, Washington DC.
- CIMTU, A. (2010). *Mobilité des personnes - Enquête Origine-Destination 2008*. Montréal, Québec: Centre d'Information Métropolitain sur le Transport Urbain, Agence Métropolitaine de Transport.
- Cirillo, C., Cornélis, E., Hubert, J.-P., & Toint, P. (2004). Les Enquêtes sur les Comportements de Mobilité, et Après ? *Reflets et Perspectives de la Vie Économique*, 4, 111-121.
- Cortés, C. E., Gibson, J., Gschwender, A., Munizaga, M., & Zuniga, M. (2010). *Bus Commercial Speed Diagnosis Based on GPS Monitored Data*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.

- Crainic, T. G., & Florian, M. (2005). Modèles d'optimisation pour la planification des systèmes de transport.
- Currie, G., & Loader, C. (2010). *Bus Network Planning for Transfers and the Network Effect in Melbourne, Australia*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- CUTS. (2010). Transit Demand Analysis. Milwaukee, Wisconsin: Center for Urban Transportation Studies, University of Wisconsin.
- Dalil Essakali, M., & Chapleau, R. (2000). *Quelle Justice dans le Système de Transport de la Région de Montréal: une Question de Méthodologie?* Paper presented at the 35e Congrès Annuel de l'AQTR, Québec.
- De Grange, L., Fernández, E., & De Cea, J. (2010). A consolidated model of trip distribution. *Transportation Research Part E*, 46(1), 61-75.
- De Grange, L., Ibeas, A., & González, F. (2011). A Hierarchical Gravity Model with Spatial Correlation: Mathematical Formulation and Parameter Estimation. *Networks and Spatial Economics*, 11(3), 439-463. doi: 10.1007/s11067-008-9097-0
- De Grange, L., Troncoso, R., Ibeas, A., & González, F. (2009). Gravity Model Estimation with Proxy Variables and the Impact of Endogeneity on Transportation Planning. *Transportation Research Part A*, 43, 105-116.
- Dessouky, M., Hall, R., Nowroozi, A., & Mourikas, K. (1999). Bus Dispatching at Timed Transfer Transit Stations Using Bus Tracking Technology. *Transportation Research Part C*, 7(4), 187-209.
- Eash, R. (1983). Several More Improvements in Understanding, Calibrating, and Applying the Opportunity Model. Chicago, IL: Chicago Area Transportation Study.
- Eash, R. (1984). Development of a Doubly Constrained Intervening Opportunity Model for Trip Distribution. Chicago, IL: Chicago Area Transportation Study.
- ECMT. (1999). Research on Transport Economics (Vol. 10). Paris: OECD.
- Evans, N. J., & Pooler, J. (1987). Distance Deterrence Effects in Constrained Spatial Interaction Models of Interprovincial Migration. *Canadian Journal of Regional Science*, Autumn 1987, 259-279.

- Faouzi, N.-E. E., Leung, H., & Kurian, A. (2011). Data fusion in intelligent transportation systems: Progress and challenges – A survey. *Information Fusion*, 12(1), 4-10. doi: <http://dx.doi.org/10.1016/j.inffus.2010.06.001>
- Fedra, K. (2003). Sustainable Urban Transportation: A Model-Based Approach. <http://www.ess.co.at/FTP/sutra.pdf>
- Feng, S., & Li, X. (2004). *Research on Public Transit Passenger OD Matrix Estimation*. Paper presented at the 8th International Conference on Applications of Advanced Technologies in Transportaion Engineering, Beijing, China.
- Feng, T., & Timmermans, H. J. P. (2014). *Comparison of Advanced GPS Data Imputation Algorithms for Detection of Transportation Mode and Activity Episode*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Ferdous, N., Bhat, C., Vana, L., Schmitt, D., Bowman, J., Bradley, M., & Pendyala, R. (2011). Comparison of Four-Step Versus Tour-Based Models in Predicting Travel Behavior Before and After Transportation System Changes – Results Interpretation and Recommendations: Ohio Department of Transportation.
- Flyvbjerg, B., Holm, M. K. S., & Buhl, S. L. (2006). Inaccuracy in Traffic Forecasts. *Transport Reviews*, 26(1), 1-24.
- Fu, L., Liu, Q., & Calamai, P. (2003). *Real-time Optimization Model for Dynamic Scheduling of Transit Operations*. Paper presented at the 82nd Annual Meeting of the Transportation Research Board, Washington DC.
- Furth, P. G., Hemily, B., Muller, T. H. J., & Strathman, J. G. (2006). Using Archived AVL-APC Data to Improve Transit Performance and Management. Washington, D.C.: TCRP.
- Galvan, P. A. F. (1995). *Une Approche Économétrique au Problème Distribution-Affectation*. (PhD), Université de Montréal, Montréal, Canada.
- Gandavarapu, S. (2012). *Using Google Transit Feed Specification in Travel Modeling*. Paper presented at the 4th Transportation Research Board Conference on Innovations in Travel Modeling, Tampa, Florida.
- Godin, A. (2012). *L'accessibilité en transport : méthodes et indicateurs*. (Maîtrise en sciences appliquées), École Polytechnique de Montréal, Montréal, Québec.

- Golob, T. F., & Hensher, D. A. (2007). The Trip Chaining Activity of Sydney Residents: A Cross-section Assessment by Age Group with a Focus on Seniors. *Journal of Transport Geography*, 15, 298-312.
- Goncalves, M. B., & De Cursi, J. E. S. (2001). Parameter Estimation in a Trip Distribution Model by Random Perturbation of a Descent Method. *Transportation Research Part B*, 35, 137-161.
- Goncalves, M. B., & Ulysséa Neto, I. (1993). The Development of a new Gravity - Opportunity Model for Trip Distribution. *Environment and Planning A*, 25(6), 817-826.
- Google. (2012). What is GTFS? Retrieved July 7, 2012, from <https://developers.google.com/transit/gtfs/>
- Goulias, K. G., Ravulaparthi, S. K., Konduri, K. C., & Pendyala, R. M. (2014). *Using Synthetic Population Generation to Replace Sample and Expansion Weights in Household Surveys for Small Area Estimation of Population Parameters*. Paper presented at the Transportation Research Board Annual Meeting, Washington, DC.
- Guihaire, V., & Hao, J.-K. (2008). Transit Network Design and Scheduling: A Global Review. *Transportation Research Part A*, 42, 1251-1273.
- Guo, Z., & Wilson, N. H. M. (2010). *Transfer Behavior and Path Choice in Public Transport Systems: A Case of the London Underground*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- Hadas, Y., & Ceder, A. A. (2008). Public Transit Simulation Model for Optimal Synchronized Transfers. *Transportation Research Record*(2063), 52-59.
- Hato, E., & Kitamura, R. (2008). *Data-Oriented Travel Behavior Analysis Based on Probe Person Systems*. Paper presented at the Innovations in Travel Demand Modeling, Austin, Texas.
- Hensher, D. A. (1977). *Urban Transport Economics*. Cambridge: Cambridge University Press.
- Hensher, D. A., & Button, K. J. (2008). *Handbook of Transport Modelling*. Amsterdam, The Netherlands: Elsevier.
- Herrera, J. C., B., W. D., Herring, R., J., B. X., Q., J., & M., B. A. (2010). Evaluation of Traffic Data Obtained Via GPS-Enabled Mobile Phones: The Mobile Century Field Experiment. *Transportation Research Part C: Emerging Technologies*, 18(4), 568-583.

- Hossein Rashidi, T., & Hasegawa, H. (2014). *An Innovative Simultaneous System of Disaggregate Models for Trip Generation, Mode, and Destination Choice*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington DC.
- Hu, P., & Pooler, J. (2002). An Empirical Test of the Competing Destinations Model. *Journal of Geographical Systems*, 4, 301-323.
- Hunt, J. G. (1990). A Logit Model of Public Transport Route Choice. *ITE*, 26-30.
- Hutchinson, B. G. (1974). *Principles of Urban Transport Systems Planning*: McGraw-Hill Inc.
- IBM. (2013). *Big Data and Analytics in Travel and Transportation*. Somers, NY: IBM.
- Iseki, H., Taylor, B. D., & Miller, M. (2006). *The Effects of Out-of-Vehicle Time on Travel Behavior: Implications for Transit Transfers*. Sacramento, California: California Department of Transportation, Division of Research and Innovation.
- Jang, W. (2010). *Travel Time and Transfer Analysis Using Transport Smart Card Data*. Paper presented at the 89th Annual Meeting of the Transportation Research Board, Washington DC.
- Javanmardi, M., Rashidi, T. H., & Mohammadian, A. K. (2010). *A Household Travel Data Simulation Tool: Software and its Applications for Impact Analysis*. Paper presented at the TRB 2010 Annual Meeting, Washington, D. C.
- Jin, P. J., Cebelak, M., Yang, F., Ran, B., & Walton, C. M. (2014). *Location-Based Social Networking Data: An Exploration into the Use of a Doubly-Constrained Gravity Model for Origin-Destination Estimation*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Kato, H., Inagi, A., & Thet Htun, P. (2010). Route Choice Analysis of Bus-use Commuters in Yangon, Myanmar. [http://www.trip.t.u-tokyo.ac.jp/kato/WP/2009/2009wp\\_e9.pdf](http://www.trip.t.u-tokyo.ac.jp/kato/WP/2009/2009wp_e9.pdf)
- Kennedy, P. E. (2003). *A Guide to Econometrics*. Bodmin, The United Kingdom: MPG Books.
- Kermanshah, M. (2004). *Notes de Cours "Prévision de la Demande en Transport"*. Sharif University of Technology. Tehran.
- Kieu, L.-M., Bhaskar, A., & Chung, E. (2014). *Transit Passenger Segmentation Using Travel Regularity Mined from Smart Card Transactions Data*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Knudsen, D. C., & Fotheringham, A. S. (1986). Matrix Comparison, Goodness-of-Fit and Spatial Interaction Modeling. *International Regional Science Review*, 10(2), 127-147.

- Komanduri, A., & Proussaloglou, K. (2010). *Getting the Most out of a Transit Onboard Survey: Results from Madison*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- Kompil, M., & Celik, H. M. (2013). Modelling trip distribution with fuzzy and genetic fuzzy systems. *Transportation Planning and Technology*, 36(2), 170-200.
- Krygsman, S., Dijst, M., & Arentze, T. (2004). Multimodal Public Transport: An Analysis of Travel Time Elements and the Interconnectivity Ratio. *Transport Policy*, 11, 265-275.
- Larousse. (2014). Contenu en ligne *Encyclopédie Larousse*.
- Le Gallou, A. (2009). *Modèle d'Affectation Pluraliste de Passagers dans un Réseau de Transport en Commun*. (MSc), École Polytechnique de Montréal, Montréal, Canada.
- Leurent, F. (2007). *Modélisation de la Demande de Transport - Généralités*. Paris, France: ENPC, Laboratoire Ville Mobilité Transport.
- Liao, C.-F., & X. Liu, H. (2010). *Development of a Data Processing Framework for Transit Performance Analysis*. Paper presented at the 89th Annual Meeting of the Transportation Research Board, Washington, DC.
- Malinovskiy, Y., Wu, Y.-J., Wang, Y., & Lee, U.-K. (2010). *Field Experiments on Bluetooth-based Travel Time Data Collection*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- Mandelzys, M., & Hellinga, B. (2010). *Automatically Identifying the Causes of Bus Transit Schedule Adherence Performance Issues using AVL/APC Archived Data*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- Manheim, M. L. (1979). *Fundamentals of Transportation Systems Analysis: Basic Concepts*: MIT Press.
- Maoh, H., Kanaroglou, P., Scott, D. M., Paez, A., & Newbold, K. B. (2009). IMPACT: An Integrated GIS-based Model for Simulating the Consequences of Demographic Changes and Population Ageing on Transportation. *Computers, Environment and Urban Systems*, 33, 200-210.
- Martel Poliquin, E. (2012). *Mieux comprendre les déterminants du choix modal*. (Maîtrise en sciences appliquées), Ecole Polytechnique, Montreal, Montréal, Québec.



- McCord, M. R., Mishalani, R. G., Goel, P., & Strohl, B. (2010). *Empirical Comparative Assessment of the IPF Procedure for Determining Bus Route Passenger OD Flows*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- McFadden, D. (1974). *Conditional Logit Analysis of Qualitative Choice Behavior*. New York: Academic Press.
- McNally, M. G., & Rindt, C. R. (2007). *The Activity-Based Approach*. Irvine, California: Center for Activity Systems Analysis, Institute of Transportation Studies, University of California.
- Mercado, R., & Paez, A. (2009). Determinants of Distance Traveled with a Focus on the Elderly: A Multilevel Analysis in the Hamilton CMA, Canada. *Journal of Transport Geography*, 17, 65-76.
- Merseytravel. (2001). *Swift card Trial Formy - Report on Focus Group and Quantitative Research*. Transport and Travel Research.
- Meyer, M. D., & Miller, E. J. (2001). *Urban Transportation Planning, A Decision Oriented Approach*: McGraw-Hill.
- Mishra, S., Wang, Y., Zhu, X., Moeckel, R., & Mahaparta, S. (2013). *Comparison between Gravity and Destination Choice Models for Trip Distribution in Maryland*. Paper presented at the TRB of the National Academies of Science Annual Meeting, Washington, DC.
- Montini, L., Rieser-Schüssler, N., Horni, A., & Axhausen, K. W. (2014). *Trip Purpose Identification from GPS Tracks*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Morency, C., & Trépanier, M. (1998). Avant-propos: les enquêtes montréalaises avant 1970. Retrieved 22 novembre 2010, 2010, from <http://www.transport.polymtl.ca/eodmtl/historiq.htm>
- Morency, C., Trépanier, M., & Agard, B. (2006a). *Analysing the variability of transit users behaviour with smart card data*. Paper presented at the IEEE Intelligent Transportation Systems, Toronto, Canada.
- Morency, C., Trépanier, M., & Agard, B. (2006b). *Contributions des données de cartes à puces à la compréhension des rythmes d'activités en transport*. Paper presented at the Association Québécoise du Transport et des Routes, Québec, Canada.

- Morency, C., Trépanier, M., & Agard, B. (2007). Measuring transit use variability with smart-card data. *Transport Policy*, 14, 193–203.
- Morency, C., Verreault, H., & Bourbonnais, P. L. (2013). Évaluation des potentialités du WEB comme outil de collecte de données sur la mobilité. Montréal, Québec: Chaire Mobilité pour le ministère des transports du Québec.
- Morgan, D. J. (2002). *A Microscopic Simulation Laboratory for Advanced Public Transportation System Evaluation*. (M. Sc.), Massachusetts Institute of Technology.
- Munizagaa, M. A., & Palmab, C. (2012). Estimation of a disaggregate multimodal public transport Origin–Destination matrix from passive smartcard data from Santiago, Chile. *Transportation Research Part C: Emerging Technologies*, 24, 9-18.
- Nazem, M., Trépanier, M., & Morency, C. (2011). Demographic Analysis of Route Choice for Public Transit. *Transportation Research Record*, 2(2217), 71-78.
- Nazem, M., Trépanier, M., & Morency, C. (2012). *Hierarchical Intervening Opportunities Model for Public Transit Trip Distribution*. Paper presented at the Conference on Advanced Systems for Public Transit (CASPT), Pontificia Universidad Catolica de Chile, The Ritz-Carlton, Santiago, Chile.
- Nazem, M., Trépanier, M., & Morency, C. (2013). Integrated Intervening Opportunities Model for Public Transit Trip Generation-Distribution: A Supply-dependent Approach. *Transportation Research Record*, 1(2350), 47-57. doi: 10.3141/2350-06
- Newbold, K. B., Scott, D. M., Spinney, J. E. L., & Kanaroglou, P. (2005). Travel Behavior within Canada's Older Population: A Cohort Analysis. *Journal of Transportation Geography*, 13, 340-351.
- OECD. (2013). *Big Data and Transport - Overview*. Paper presented at the International Transport Forum, Leipzig, Germany.
- Ortuzar, J. d. D., & Willumsen, L. G. (1994). *Modelling Transport*. Chichester, England: John Wiley & Sons Ltd.
- Pan, C., Lu, J., Di, S., & Ran, B. (2006). Cellular-Based Data-Extracting Method for Trip Distribution. *Transportation Research Record*(1945), 33-39.
- Park, J. Y., & Kim, D. J. (2008). *The Potential of Using the Smart Card Data to Define the Use of Public Transit in Seoul*. Paper presented at the Annual Meeting of the Transportation Research Board, Washington DC.

- Pelletier, M.-P., Trépanier, M., & Morency, C. (2009). *Smart Card Data in Public Transit Planning: A Review*. Montreal, Canada: CIRRELT.
- Pendyala, R. M., Kitamura, R., & Kikuchi, A. (2004). *FAMOS: The Florida Activity Mobility Simulator*. Paper presented at the Progress in Activity-Based Analysis, Maastricht, The Netherlands.
- Perrakis, K., Karlis, D., Cools, M., Janssens, D., Vanhoof, K., & Wets, G. (2012). A Bayesian approach for modeling origin–destination matrices. *Transportation Research Part A: Policy and Practice*, 46(1), 200-212. doi: <http://dx.doi.org/10.1016/j.tra.2011.06.005>
- Rajesson, F. (2009). *Modèle Hybride d'Estimation de la Demande de Transport Collectif*. (M.Sc.), École Polytechnique de Montréal, Montréal, Québec.
- Raveau, S., & Muñoz, J. C. (2014). *Analyzing route choice strategies on transit networks*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Richardson, A. J., Ampt, E. S., & Meyburg, A. H. (1995). *Survey Methods for Transport Planning*. Melbourne: Eucalyptus Press.
- Riegel, L., & Attanucci, J. (2014). *Utilizing Automatically Collected Smart Card Data to Enhance Travel Demand Surveys*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Rohne, A. (2014). *Building Open-Source Tools to Perform Quality Control Checks on Transit On-Board Survey Data*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Roorda, M. J., & Miller, E. J. (2006). Assessing Transportation Policy Using an Activity-Based Microsimulation Model of Travel Demand. *ITE Journal*, 76(11), 16-21.
- Roshan Zamir, K., Nasri, A., Baghaei, B., Mahapatra, S., & Zhang, L. (2014). *Comparative Analysis of the Effect of Transit-Oriented Development on Trip Generation, Distribution, and Mode Share in the Washington, D.C. and Baltimore Metropolitan Areas*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Schneider, M. (1959). Gravity Models and Trip Distribution Theory. *Regional Science Association*, 1, 51-56.
- Shan, R., Zhong, M., Du, D., & Lu, C. (2012). *Comparison of Trip Generation Results from Activity-based and Traditional Four-Step Travel Demand Modeling: A Case Study of*

- Tampa, Florida*. Paper presented at the Conference of the Transportation Association of Canada, Fredericton, NB.
- Sheffi, Y. (1985). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- Sherali, H. D., Narayanan, A., & Sivanandan, R. (2003). Estimation of Origin-Destination Trip-tables Based on a Partial Set of Traffic Link Volumes. *Transportation Research Part B*, 37, 815-836.
- Shi, X., & Lin, H. (2014). *Analysis of Bus Commuters' Travel Characteristics Using Smart Card Data: Case of Shenzhen, China*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Siddique, A. J., & Khan, A. M. (2006). Microscopic Simulation Approach to Capacity Analysis of Bus Rapid Transit Corridors. *Journal of Public Transportation*, 181-200.
- Simard, S., Springate, E., & Casello, J. M. (2010). *The Development and Deployment of GIS Tools to Facilitate Transit Network Design and Operational Evaluation*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- Simas Oliveira, M. G., & Casas, J. (2010). *Improving Data Quality, Accuracy, and Response in On-Board Surveys: An Application of Innovative Technologies*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- Simas Oliveira, M. G., Vovsha, P., Wolf, J., & Mitchell, M. (2014). *Evaluating Two Methods for Identifying Trip Purpose in GPS-Based Household Travel Surveys*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Smith, D. P., & Hutchinson, B. G. (1981). Goodness-of-Fit Statistics for Trip Distribution Models. *Transportation Research*, 15A, 295-303.
- Spurr, T., Chapleau, R., & Piché, d. (2014). *Discovery and Partial Correction of Travel Survey Bias Using Subway Smart Card Transactions*. Paper presented at the Transportation Research Board Annual Meeting, Washington, DC.
- StatCan. (2010). Population Projections for Canada, Provinces and Territories: Statistics Canada.
- StatCan. (2011a). 2006 Census of Population, Quebec - Cat. No. 94-581-X2006002. <http://ivt.crepuq.qc.ca/> [Accessed October 11, 2011]
- StatCan. (2011b). Statistics Canada. [www.statcan.gc.ca](http://www.statcan.gc.ca)

- Stouffer, S. A. (1940). Intervening Opportunities: A Theory Relating Mobility and Distance. *American Sociological Review*, 5(6), 845-867.
- Stouffer, S. A. (1960). Intervening Opportunities and Competing Migrants. *Journal of Regional Science*, 2, 1-26.
- Tamin, O. Z., & Sulistyorini, R. (1997). Public Transport Demand Estimation by Calibrating the Combined Trip Distribution-Mode Choice (TDMC) Model from Passenger Counts. *Journal of Advanced Transportation*, 31(1), 5-18.
- Thamizh Arasan, V., Wermuth, M., & Srinivas, B. S. (1996). Modelling of stratified urban trip distribution. *Journal of Transportation Engineering*, 122(5), 342-349.
- Thomas, T., & Tutert, S. I. A. (2013). An empirical model for trip distribution of commuters in The Netherlands: transferability in time and space reconsidered. *Journal of Transport Geography*, 26(0), 158-165. doi: <http://dx.doi.org/10.1016/j.jtrangeo.2012.09.005>
- Thévenin, T. (2001). L'analyse désagrégée : Un Luxe ou une Nécessité ? Application à l'Analyse des Générateurs de Déplacements à l'Échelle Urbaine. <http://thema.univ-fcomte.fr/theoq/pdf/2001/thevenin.pdf> [Consulté le 20 novembre 2009]
- Toronto, U. o. (2008). Transportation Tomorrow Survey 2006 - Data Validation. Toronto, Ontario: Transportation Information Steering Committee, University of Toronto.
- Train, K. E. (2009). *Discrete Choice Methods with Simulation*: Cambridge University Press.
- Tranchant, N. (2005). *Modèle de Dérivation des Déplacements en Transport Collectif à partir de Données de Cartes à Puces*. (M. Sc. A.), École Polytechnique, Montréal, Canada.
- Trépanier, M. (1999). *Modélisation Totale Désagrégée et Orientée-Objet Appliquée aux Transports Urbains*. (PhD), École Polytechnique de Montréal, Montréal.
- Trépanier, M., & Chapleau, R. (1996). *Un Modèle d'Analyse Désagrégée des Générateurs de Déplacements: Vers une Connaissance Détaillée de l'Utilisation des Lieux Urbains*. Paper presented at the 31e Congrès de l'Association Québécoise du Transport et des Routes, Mont Sainte-Anne.
- Trépanier, M., & Chapleau, R. (2001). Analyse Orientée-Objet et Totale Désagrégée des Données d'Enquêtes Ménages Origine-Destination. *Canadian Journal of Civil Engineering*, 28, 48-58.

- Trépanier, M., Chapleau, R., & Morency, C. (2008). Tools and Methods for a Transportation Household Survey. *Journal of the Urban and Regional Information Systems Association*, 20(1). [https://www.urisa.org/files/\\_complete.pdf#page=37](https://www.urisa.org/files/_complete.pdf#page=37)
- Trépanier, M., Morency, C., & Agard, B. (2009). Calculation of Transit Performance Measures Using Smartcard Data. *Journal of Public Transportation*, 12(1), 79-96.
- Trépanier, M., Morency, C., & Blanchette, C. (2008). *Les Systèmes de Paiement par Cartes à Puces: un Complément aux Enquêtes Origine-Destination?* Paper presented at the 43ème Congrès de l'Association Québécoise du Transport et des Routes, Québec.
- Trépanier, M., Morency, C., Chapleau, R., & Blanchette, C. (2008). *Fusion of Smart Card Data and Household Travel Survey Data for the Enhancement of Transit Riders' Behavior Characterisation*. Paper presented at the International Conference on Survey Methods in Transport: Harmonisation and data comparability, Annecy, France.
- Trépanier, M., Morency, C., & Rajesson, F. (2008). Design de Réseau de Transport, Rapport d'Avancement, Partie Estimation de la Demande. Montréal: École Polytechnique.
- Trépanier, M., Tranchant, N., & Chapleau, R. (2007). Individual Trip Destination Estimation in a Transit Smart Card Automated Fare Collection System. *Journal of Intelligent Transportation Systems - Technology, Planning, and Operations*, 11(1), 1-14.
- Uniman, D. L., Attanucci, J., Mishalani, R. G., & Wilson, N. (2010). *Service Reliability Measurement using Automated Fare Card Data: Application to the London Underground*. Paper presented at the 89th Annual Meeting of the Transportation Research Board, Washington DC.
- Utsunomiya, M., Attanucci, J., & Wilson, N. H. (2006). Potential Uses of Transit Smart Card Registration and Transaction Data to Improve Transit Planning. *Transportation Research Record*(1971), 119-126.
- Varley, D. M., & Chen, C. (2010). *If You Build it, They will Come? A Dynamic Analysis of the Factors Affecting the Transit Ridership in the New York City Region*. Paper presented at the TRB 2010 Annual Meeting, Washington, D.C.
- Veenstra, S. A., Thomas, T., & Tutert, S. I. A. (2010). *Trip distribution for limited destinations: A case study for grocery shopping trips in the Netherlands*. Paper presented at the 89th Annual Meeting of the Transportation Research Board, Washington DC.

- Viggiano, C., Koutsopoulos, H. N., & Attanucci, J. (2014). *User Behavior in Multi-Route Bus Corridors: Analysis via a Web-Based Survey*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Volosin, S. E., Pendyala, R. M., Kerrigan, J., Greene, E., Livshits, V., & Samuelson, J. (2014). *Measuring Travel Characteristics of a University Population: Experiences from Design and Administration of Web-Based Travel Survey*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Vrtic, M., Frohlich, P., Schussler, N., Axhausen, K. W., Lohse, D., Schiller, C., & Teichert, H. (2007). Two-dimensionally Constrained Disaggregate Trip Generation, Distribution and Mode Choice model: Theory and Application for a Swiss National Model. *Transportation Research Part A*, 41, 857-873.
- Wang, J., Chen, S., He, Y., & Gao, L. (2006). Simulation of Transfer Organization of Urban Public Transportation Hubs. *Journal of Transportation Systems Engineering and Information Technology*, 6(6), 96-102.
- Webster, F. V., & Bly, P. H. (1982). The Demand for Public Transport. Part II. Supply and Demand Factors of Public Transport. *Transport Reviews*, 2(1), 23-46.
- Wikipédia. (2014). Contenu en ligne *Encyclopédie Wikipédia*.
- Wills, M. J. (1986). A flexible gravity-opportunities model for trip distribution. *Transportation Research Part B: Methodological*, 20(2), 89-111.
- Wilson, A. G. (1970). Advances and Problems in Distribution Modelling. *Transportation Research*, 4, 1-18.
- Woldeamanuel, M., & Covington, K. (2014). *Public Transit Spatial Mismatch in Los Angeles Suburbs: Socio-spatial Distribution of Service Quality, Neighborhood Need, and Labor Force Participation*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Wong, J. C. (2013). *Use of the General Transit Feed Specifications (GTFS) in Transit Performance Measurement*. (Master of Science in Civil Engineering), Georgia Institute of Technology.
- Xu, M., Chen, A., & Gao, Z. (2008). An Improved Origin-Based Algorithm for Solving the Combined Distribution and Assignment Problem. *European Journal of Operational Research*, 188, 354-369.

- Yaldi, G., Taylor, M. A. P., & Yue, W. L. (2011). *Forecasting Origin-Destination Matrices by Using Neural Network Approach: A Comparison of Testing Performance between Back Propagation, Variable Learning Rate and Levenberg-Marquardt Algorithms*. Paper presented at the Australasian Transport Research Forum 2011 Proceedings, Adelaide, Australia.
- Yun, M.-P., & Liu, X.-W. (2014). *Commuter Mode Choice Forecast Considering Bus Rapid Transit under Planning: A Case Study of Yichang, China*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington, DC.
- Zakeri Sohi, M. (2007). *Iran Network and Transit Modeling and Forecasting Using EMME/2*. Paper presented at the 21st International EMME Users Conference, Toronto.
- Zhou, Z., Chen, A., & Wong, S. C. (2009). Alternative Formulations of a Combined Trip Generation, Trip Distribution, Modal Split, and Trip Assignment Model. *European Journal of Operational Research*, 198, 129-138.
- Zhu, Y., & Ferreira Jr., J. (2014). *Synthetic Population Generation at Disaggregated Spatial Scales for Land Use and Transportation Microsimulation*. Paper presented at the Transportation Research Board 93rd Annual Meeting, Washington DC.



## ANNEXES

### A. SYNTHÈSE DE LA REVUE DE LITTÉRATURE

Tableau A.1 Synthèse chronologique de la revue de littérature

Thèmes	Chercheurs	Sujet
Généralités de la modélisation des déplacements en transport	(Manheim, 1979)	Principes généraux de la modélisation en transport
	(Meyer & Miller, 2001; Ortuzar & Willumsen, 1994)	PSC
	(McNally & Rindt, 2007; Roorda & Miller, 2006)	Modèles basés sur les activités
<b>APPROCHE AGRÉGÉE</b>		
Généralité de la PSC et modèles de génération des déplacements	(Bonnell, 2004)	PSC
	(Hensher & Button, 2008)	PSC avec rétroaction et PSC sans rétroaction
	(Boyce et al., 2008)	Différentes méthodes pour l'application de la rétroaction dans la PSC
	(Rajesson, 2009)	Méthode hybride de la PSC
	(De Grange et al., 2009)	Modèle mathématique d'optimisation pour la distribution des déplacements en utilisant des destinations compétitives
	(Ferdous et al., 2011)	Comparaison de la PSC et les modèles basés sur les chaînes de déplacements
	(Woldeamanuel & Covington, 2014)	Déterminant de demande de transport en lien avec des attributs de la population et le niveau de service en TC

Tableau A.1 Synthèse chronologique de la revue de littérature (suite)

Thèmes	Chercheurs	Sujet
Distribution des déplacements	(Stouffer, 1940, 1960)	Idée initiale du modèle d'opportunités
	(Schneider, 1959)	Modèle d'opportunités dans sa façon actuelle
	(Hensher, 1977; Ortuzar & Willumsen, 1994; Wilson, 1970)	Modèles de distributions de déplacements
	(Eash, 1983, 1984)	Application du modèle d'opportunités dans la région de Chicago, Illinois
	(Thamizh Arasan et al., 1996)	Indicateurs de performance pour le modèle gravitaire
	(Bonnell, 2004)	Formalisations classiques des modèles de distribution
	(Kermanshah, 2004)	Base comportementale du modèle d'opportunités
	(S. Feng & Li, 2004)	Calibration des modèles de distribution avec les données provenant de comptage de passagers aux arrêts
	(De Grange et al., 2009)	Modèle de distribution développé à partir des notions d'accessibilité et d'attractivité
	(Caceres et al., 2007; Herrera et al., 2010; Pan et al., 2006)	Utilisation des données de cellulaires pour le développement d'une matrice de référence OD
	(Blogg et al., 2010)	Utilisation des données de Bluetooth pour le développement d'une matrice de référence OD
	(Munizagaa & Palmab, 2012)	Utilisation des données de CAP pour le développement d'une matrice de référence OD

Tableau A.1 Synthèse chronologique de la revue de littérature (suite)

Thèmes	Chercheurs	Sujet
	(Celik, 2010)	Taille d'échantillon requise pour le développement d'un modèle gravitaire afin de produire des mesures globales
	(De Grange et al., 2010)	Approche consolidée pour la distribution des déplacements et la formulation d'optimisation multi-objective pour déterminer la matrice OD
	(Cheng, Wilmot, & Baker, 2011)	Modèle gravitaire dynamique pour étudier le choix de destination lors d'une catastrophe naturelle
	(De Grange, Ibeas, & González, 2011)	Modèle gravitaire hiérarchique pour considérer certaines corrélations spatiales dans le processus de la modélisation
	(Afandizadeh & Hamedani, 2012)	Modèle d'opportunités avec une approche floue pour le motif de magasinage
	(Perrakis et al., 2012)	Approche Bayes pour le développement des matrices OD
	(Kompil & Celik, 2013)	Modélisation de la distribution des déplacements avec une approche floue
	(Thomas & Tutert, 2013)	Transférabilité des modèles de distribution de type gravitaire dans le temps et l'espace
	(Jin et al., 2014)	Utilisation des données provenant des médias sociaux pour le développement d'un modèle gravitaire
Répartition modale	(Bonnell, 2004)	Variables indépendantes associées au choix de modes par des individus
	(Train, 2009)	Principes des modèles de choix discrets

Tableau A.1 Synthèse chronologique de la revue de littérature (suite)

Thèmes	Chercheurs	Sujet
	(Yun & Liu, 2014)	Choix modal en TC à partir des modèles de préférences déclarée/révlée (étude de cas : bus à haut niveau de service)
Affectation des déplacements	(McFadden, 1974)	Procédure du maximum de vraisemblance pour calibrer les modèles de choix discret
	(Sheffi, 1985)	Différentes méthodes mathématiques pour l'affectation de déplacements
	(Hunt, 1990)	Choix d'itinéraires en TC (motif travail à la destination du centre-ville)
	(Ortuzar & Willumsen, 1994)	Différents modèles d'affectation des déplacements
	(Bonnel, 2004)	Principes de base des modèles d'affectation des déplacements
	(Krygsman et al., 2004)	Étude des attributs de temps de déplacement dans un contexte multimodal en TC et appui sur les modes d'accès au réseau
	(Bovy & Hoogendoorn-Lanser, 2005)	Étude des pénalités relatives aux divers attributs de temps dans un réseau multimodal de TC
	(Crainic & Florian, 2005)	Modèle d'optimisation pour l'affectation des déplacements
	(Iseki et al., 2006)	Étude des effets de différents composants de temps hors de véhicule notamment dans les correspondances sur le comportement des voyageurs
	(Hensher & Button, 2008)	Différents modèles d'affectation des déplacements

Tableau A.1 Synthèse chronologique de la revue de littérature (suite)

Thèmes	Chercheurs	Sujet
	(Kato et al., 2010)	Modèle Logit binaire pour déterminer le choix d'itinéraires dans un réseau de bus
	(Nazem et al., 2011)	Considération des caractéristiques sociodémographiques dans le choix d'itinéraires en TC
	(Raveau & Muñoz, 2014)	Analyse des stratégies de voyageurs pour réaliser leur choix d'itinéraires en TC
Intégration des étapes de la PSC	(Galvan, 1995)	Approche économétrique pour intégrer les étapes de distribution et d'affectation en TC
	(Tamin & Sulistyorini, 1997)	Modèle de distribution-répartition modale intégré calibré avec les données de comptes à bord
	(Xu et al., 2008)	Modèle de distribution-affectation intégré
	(Rajesson, 2009)	Modèle hybride pour estimer la demande en TC
	(Zhou et al., 2009)	Modèle totalement intégré calibré avec une programmation mathématique
	(Hossein Rashidi & Hasegawa, 2014)	Intégration des étapes de la PSC dans l'optique du développement d'un modèle désagrégé
Intégration de la PSC avec des modèles démographiques	(Maoh et al., 2009)	Modèle IMPACT (démographie - transport - environnement) intégré
	(Roshan Zamir, Nasri, Baghaei, Mahapatra, & Zhang, 2014)	Impact des interventions d'urbanistiques sur les étapes de la PSC

Tableau A.1 Synthèse chronologique de la revue de littérature (suite)

Thèmes	Chercheurs	Sujet
<b>APPROCHE DÉSAGRÉGÉE</b>		
Modèles désagrégés	(Bonnell, 2004)	Différences entre les démarches agrégée et désagrégée en modélisation
	(Vrtic et al., 2007)	Utilisation des modèles mathématiques dans l'approche désagrégée de génération - distribution - répartition modale intégrée
	(Babazadeh et al., 2010)	Développement des modèles désagrégés par le biais des modèles d'optimisation mathématiques
<b>APPROCHE DESCRIPTIVE</b>		
Généralités	(ECMT, 1999)	Définition générale de l'approche descriptive en planification de transport
	(Cirillo et al., 2004)	Approche descriptive comme étape en aval des enquêtes OD
Sources de données	(Bagchi & White, 2004)	Différentes technologies et utilisations des données de CAP
	(Morency et al., 2006a, 2006b, 2007; Park & Kim, 2008; Utsunomiya et al., 2006)	Utilisation des données de CAP pour la planification en TC
	(Trépanier, Morency, & Blanchette, 2008)	Enquête OD
	(Rajesson, 2009; Trépanier, Morency, & Rajesson, 2008)	Classification des types de données en transport
	(Pelletier et al., 2009)	Utilisation des données de CAP dans la planification en TC

Tableau A.1 Synthèse chronologique de la revue de littérature (suite)

<b>Thèmes</b>	<b>Chercheurs</b>	<b>Sujet</b>
	(Chu & Chapleau, 2010)	Enrichissement des données de CAP
	(Malinovskiy et al., 2010)	Utilisation des données de Bluetooth pour calculer le temps de parcours en TC
	(IBM, 2013; OECD, 2013)	Potentiel de big data en TC
	(Bonnell et al., 2014)	Utilisation des données de cellulaires pour le développement des matrices OD
Enquêtes OD, MADITUC et approche orientée-objet	(Chapleau, 1986)	MADITUC
	(Richardson et al., 1995)	Niveaux d'informations dans les enquêtes OD
	(Chapleau & Lavigneur, 1996)	MADITUC pour étudier les caractéristiques associées aux lieux d'emplois et aux travailleurs
	(Trépanier & Chapleau, 1996)	MADGEN
	(Chapleau & Trépanier, 1997; Trépanier, 1999)	MOOT
	(Chapleau et al., 1998)	Présentation du passage graduel des niveaux d'information en transport, de haut à fin (approche floue)
	(Dalil Essakali & Chapleau, 2000)	MADITUC pour faire l'étude des effets redistributifs et les problématiques fiscales
	(Chapleau et al., 2001; CIMTU, 2010; Cirillo et al., 2004)	Enquête OD montréalaise
	(Thévenin, 2001)	Analyse désagrégée en transport

Tableau A.1 Synthèse chronologique de la revue de littérature (suite)

Thèmes	Chercheurs	Sujet
	(Trépanier & Chapleau, 2001)	Modèle orienté-objet enrichi en transport
	(Chapleau, 2002)	Utilisation du MADITUC en planification de TC
	(Morency et al., 2007)	Utilisation de la MOOT pour analyser la variabilité d'usage de TC à l'aide des données de CAP
	(Trépanier, Chapleau, et al., 2008)	MADITUC et approche orientée-objet en transport
	(Morency et al., 2013; Viggiano et al., 2014; Volosin et al., 2014)	Enquêtes web
	(Spurr et al., 2014)	Biais des enquêtes OD
Enquête à bord	(Currie & Loader, 2010)	Analyse descriptive des effets de correspondances sur le comportement des voyageurs dans un réseau de TC
	(Komanduri & Proussaloglou, 2010; Simas Oliveira & Casas, 2010)	Stratégies d'amélioration des enquêtes à bord
	(X. Chen et al., 2010)	Analyse de fiabilité d'un réseau de bus à l'aide des données d'enquêtes à bord
	(Rohne, 2014)	Validation des enquêtes à bord par le biais d'un outil développé dans ArcMap d'ESRI
Données opérationnelles et passives	(Berkow et al., 2009; Bertini & El-Geneidy, 2003)	Étude de performance d'un système de TC à l'aide de données archivées



Tableau A.1 Synthèse chronologique de la revue de littérature (suite)

Thèmes	Chercheurs	Sujet
	(Furth et al., 2006; Mandelzys & Hellinga, 2010)	Étude de ponctualité de services de TC à l'aide des données d'AVL et d'APC
	(Morency et al., 2006a, 2006b, 2007)	Utilisation des données de CAP pour l'analyse de comportements des voyageurs
	(Park & Kim, 2008; Utsunomiya et al., 2006)	Utilisation des données de CAP dans la planification en TC
	(Trépanier et al., 2007)	Méthode pour dériver les destinations à partir des données de CAP dans un système de simple-validation
	(Trépanier et al., 2009)	Calcul des indicateurs de performance en TC à l'aide des données de CAP
	(Mandelzys & Hellinga, 2010)	Utilisation des données opérationnelles et passives pour étudier la ponctualité du service de TC
	(Jang, 2010)	Étude de temps de parcours et des correspondances à l'aide des données de CAP
	(Cortés et al., 2010)	Utilisation de données de GPS pour des fins d'analyse en TC
	(Guo & Wilson, 2010; Jang, 2010)	Analyse de comportement des usagers de TC dans les correspondances, à l'aide des données de CAP dans un système de double-validations

Tableau A.1 Synthèse chronologique de la revue de littérature (suite)

Thèmes	Chercheurs	Sujet
	(McCord et al., 2010)	Dérivation de la matrice OD avec les données intégrées d'APC et d'enquête à bord
	(Uniman et al., 2010)	Étude de fiabilité d'un système de TC à l'aide des données de CAP
	(Blum et al., 2010)	Utilisation des données de compte à bord et d'une enquête OD pour la détermination des matrices OD
	(Cortés et al., 2010)	Utilisation des données de localisation des véhicules pour calculer la vitesse commerciale en TC
	(Gandavarapu, 2012; Godin, 2012; Martel Poliquin, 2012; Wong, 2013)	Utilisation des données de services planifiés (GTFS) pour la caractérisation de l'offre de TC
	(Arifn & Axhausen, 2012; Montini et al., 2014)	Enrichissement des données provenant de GPS
	(T. Feng & Timmermans, 2014; Simas Oliveira, Vovsha, Wolf, & Mitchell, 2014)	Utilisation des données de GPS pour dériver des attributs des déplacements (mode, etc.)
	(Riegel & Attanucci, 2014)	Amélioration des données d'enquêtes OD à l'aide des données passives
	(Chu, 2014; Shi & Lin, 2014)	Utilisation des données passives de CAP pour la caractérisation de la demande en TC

Tableau A.1 Synthèse chronologique de la revue de littérature (suite)

Thèmes	Chercheurs	Sujet
Systèmes d'aide à la planification	(Chapleau & Morency, 2005)	Utilisation des SIG dans les études de transport
	(Chaudhari & Ye, 2010)	Utilisation des SIG pour faire des sketches plans afin améliorer l'efficacité d'un système de bus scolaire
	(Simard et al., 2010)	Utilisation des SIG pour le design et l'analyse d'un réseau de TC
<b>APPROCHE DE SIMULATION</b>		
Simulation des données de références	(Pendyala et al., 2004)	Simulateur d'activités et de mobilité en Floride, FAMOS (déplacements obligatoires)
	(Javanmardi et al., 2010)	Simulateur de mobilité des ménages (déplacements de tout motif)
	(Goulias, Ravulaparthi, Konduri, & Pendyala, 2014)	Améliorations relatives à la pondération des enquêtes OD à l'aide de la simulation des données de référence
Simulation des scénarios en TC	(Morgan, 2002)	Évaluation des impacts des systèmes avancés en TC tels que les priorités aux véhicules de TC dans les intersections
	(Bachelet & Yon, 2005)	Méthode d'optimisation - simulation hybride
	(Siddique & Khan, 2006)	Étude des corridors exclusifs pour le bus à haute niveau de service à l'aide de la simulation
	(Wang et al., 2006)	Utilisation de la simulation pour étudier les aspects physiques dans les correspondances

Tableau A.1 Synthèse chronologique de la revue de littérature (suite)

Thèmes	Chercheurs	Sujet
	(Hadas & Ceder, 2008)	Utilisation de la simulation pour étudier l'efficacité des politiques de synchronisation dans les correspondances
	(Zhu & Ferreira Jr., 2014)	Utilisation de la simulation pour étudier l'utilisation du sol et la modélisation en transport
<b>APPROCHE OPÉRATIONNELLE</b>		
POTC	(Dessouky et al., 1999)	Dispatching des bus
	(Fu et al., 2003)	Modèle d'optimisation à temps réel pour la planification dynamique en TC
	(Ceder, 2007)	POTC
	(Guihaire & Hao, 2008)	Revue sur la POTC
	(Pelletier et al., 2009)	Niveaux d'étude en TC